Molecular evolution and genetic control of flowering in the *Eucalyptus globulus* species complex

Rebecca C. Jones BSc (Hons)

Submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy

School of Plant Science, University of Tasmania

November 2009

DECLARATION

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in the thesis. To the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright.

This thesis may be made available for loan and limited copying in accordance with the Copyright Act 1968.

Rebecca Jones

November 2009

ABSTRACT

Eucalyptus globulus (blue gum) is one of the most widely planted eucalypts in temperate parts of the world, including its own natural range in south-eastern Australia. Australia is therefore a custodian of an internationally important genetic resource and it is necessary to develop population-based strategies to manage this gene pool. In addition, an understanding of the flowering process in *E. globulus* has key implications in managing seed orchards, developing early-flowering trees to accelerate breeding programs, generating sterile genetically modified trees, and determining the risk of genetic pollution. This thesis consists of, firstly, a population genetic analysis of the *E. globulus* species complex and, secondly, an analysis of the flowering process in *E. globulus*, including quantitative genetics and gene expression analyses.

Eucalyptus globulus is a species complex consisting of four taxa variously described as species or subspecies (*bicostata*, *globulus*, *maidenii* and *pseudoglobulus*). The cores of these taxa are geographically and morphologically distinct, but are linked by intergrade populations that are intermediate in morphology. The intergrade populations represent a significant proportion of the distribution of the species complex, but their diverse and intermediate morphologies confound taxonomic classification. To assess the genetic affinities within the complex, nine nuclear DNA markers (microsatellites) were used to genotype 1198 trees representing 33 morphological core and intergrade populations from across the natural range of the species complex in south-eastern Australia. Combined with morphological measurements, this analysis provided insights into some of the evolutionary processes that have shaped the patterns of genetic variation in the *E. globulus* gene pool, including drift and morphological convergence, and indicated that the patterns of variation in the complex are probably due to primary differentiation rather than recent secondary contact. The analysis also identified the possible geographic origin of the species complex, evaluated the conservation status of small, isolated populations and provided a framework to guide the development of seed transfer guidelines and identify the potential impact of gene flow from *E. globulus* planted within the natural range of the complex.

Variation in flowering season is clearly a significant barrier to gene flow within the *E. globulus* complex. Surveys of *E. globulus* ssp. *globulus* provenances grown in a common environment near Hobart (Tasmania), conducted for the last six years, showed that flower bud opening occurred over a nine month period and was under strong genetic control ($H^2 = 0.69$) with highly significant differences in flower opening time among provenances. Surprisingly though, provenances did not vary in the timing of macroscopic appearance of flower buds, which occurred in October (spring) in all provenances ($H^2 = 0.06$). "Early" flower opening genotypes therefore had a shorter flower bud development time than the flower buds of "late" flower opening genotypes. The *E. globulus* homologues of *FLOWERING LOCUS T* (*FT*), *TERMINAL FLOWER1* (*TFL1*), *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1* (*SOC1*), *LEAFY* (*LFY*), *APETALA1* (*AP1*), *SHORT VEGETATIVE PHASE* (*SVP*), *SEPALLATA* (*SEP*) and *PISTILLATA* (*PI*) were isolated and quantitative RT-PCR was used to show that the expression of the homologues of *FT* and *LFY* in the leaf and apex of *E. globulus* were associated with the annual transition from vegetative to reproductive growth (i.e. flower bud initiation) over a two year period. However, there was no strong seasonal pattern of expression of *TFL1*, *SOC1*, *AP1*, *SVP*, *SEP* or *PI* in leaves, apices and/or young flower buds. In a comparison of *FT* and *LFY* expression patterns in two clones each of an early and late flowering genotype, no association between the expression of these genes and the timing of flower bud opening was shown. This indicates that *FT* and *LFY* could form part of the flower initiation pathway in *Eucalyptus* but do not regulate the observed differences in flower bud opening time.

ACKNOWLEDGEMENTS

I have many people to thank for their help with this project. I was very fortunate to have René Vaillancourt, Brad Potts, Dorothy Steane, Jim Weller and Valérie Hecht as my enthusiastic supervisors. I also gratefully acknowledge the support of the Australian Research Council as well as Timbercorp, WA Plantation Resources, Gunns Ltd, DPI Victoria, Forestry Tasmania, Forests NSW, the Southern Tree Breeding Association and Arianda Pty Ltd who were partners on this linkage grant. Financial support also came from the CRC for Forestry and School of Plant Science.

I particularly thank Martyn Lavery of Arianda Pty Ltd for his enthusiasm for the project and assistance in sampling bluegums across south-eastern Australia. James Marthick also assisted with some of the sample collection and lab preparation. I also thank the government agencies and private landholders who gave us access to trees on their land: State Forests NSW, particularly Chris Rhynehart and Ian Cotterill; NSW National Parks and Wildlife Service, especially the Merimbula office; DSE Victoria; Grant Commins at Jenolan Caves; and David McDonald and the Hill family at Mt. Bryan. Many thanks are also due to Peter Gore at seedEnergy Pty Ltd for use of the Cambridge seed orchard and who also, along with Marian McGowen, provided some of the flowering data. Thanks also to Dorothy Steane, Tim Jones and Susan Foster who provided microsatellite genotypes for the Tasmanian populations of the *E. globulus* complex. I would also like to thank Greg Dutkowski for help with mapping software; Greg Jordan, Mark Hovenden, Bob Barbour and Richard Kerr for their help with statistical analyses; and Paul Tilyard, Mick Oates, Lim Chee Liew, James Worth and Jenny Smith for other technical assistance, and especially Adam Smolenski for his help in the lab.

I'm grateful for the friendship and advice that Jaz Janes gave; she was great company for the PhD journey. Thanks also to my colleagues and friends in Plant Science for the great research and social environment that they help create.

Many thanks to my family and, of course, my friends, especially to Shan, Jo, Torz and Em for their encouragement and always scheming up lots of adventures to ensure I had a well-balanced life during my candidature. And finally thanks to Mark for his continued support, love and friendship.

TABLE OF CONTENTS

PREFACE

The *Eucalyptus globulus* (Myrtaceae) species complex is naturally distributed in south-eastern mainland Australia and on the island of Tasmania. It has a widespread distribution that is largely continuous but there are several small, disjunct populations at the periphery of its range (Brooker and Kleinig 1999) and large geographic barriers (the Bass Strait, the Great Dividing Range and the Murray Darling Depression) that are likely to be barriers to gene flow. Differences in flowering time are also expected to provide barriers to gene flow among populations. The species complex consists of four taxa variously described as species or subspecies (*bicostata*, *globulus*, *maidenii* and *pseudoglobulus*) (Kirkpatrick 1975; Pryor and Johnson 1971; Jordan *et al.* 1993; Brooker 2000) but most recently as subspecies (Slee *et al.* 2006). These taxa will hereafter be referred to by their subspecific names, hence *globulus* (etc.) refers to the subspecies and *E. globulus* refers to the species complex. The cores of these taxa are geographically and morphologically distinct, but are linked by intergrade populations that are intermediate in morphology. The origin of these intergrade populations, as well as the origin of the species complex and migration pattern out of this centre of origin, has been the subject of considerable debate (Kirkpatrick 1975, Jordan *et al.* 1993) and the use of neutral molecular marker studies can provide a valuable contribution to discussing these issues (Freeman *et al.* 2001, Steane *et al.* 2006).

Eucalyptus globulus is an important tree both ecologically and economically. As a dominant tree of south-eastern Australian lowland forests, many insects and birds, including the endangered swift parrot (*Lathamus discolor*), feed on the nectar of its flowers. The yearly reproductive success of the swift parrot is thought to depend upon the flowering intensity of *E. globulus* (Hingston *et al.* 2004a). *Eucalyptus globulus* is also the most widely planted hardwood species in temperate parts of the world, and is used mostly for pulpwood production (Eldridge *et al.* 1993). Hence, Australia is the custodian of an internationally important genetic resource and has a responsibility to develop strategies to manage this species sustainably to ensure the native gene pool of this economic resource is maintained. A gene pool management strategy will conserve the full range of genetic diversity in the species, thereby maximising its long-term evolutionary flexibility. It will also include the identification of populations of high

conservation value, determine the conservation value of outlying populations, and will help to conserve associated communities.

As well as being planted worldwide, *E. globulus* is planted in parts of south-eastern Australia, including within its own natural range. In most cases, the plantation provenance may not be local, raising concern that gene flow from these non-local plantations could affect the native gene pool of *E. globulus* (Potts *et al.* 2003). A risk assessment model is being developed at UTAS (Barbour *et al.* 2008b) to determine the probability and consequences of such gene flow, and the research presented in this thesis will contribute to this model. The development of a gene pool management strategy for *E. globulus* requires the assessment of the potential consequences of genetic contamination. The consequence of such gene flow may be related to the degree of genetic difference between the plantation provenance and the local populations. The more exotic the genetic contribution from plantations, the more likely it is to have an impact on local populations. Flowering time asynchrony is a major barrier to gene flow in plants (Levin 1978). Understanding flowering time variation in *E. globulus* will be important to assess the potential for gene flow between plantations and native forest. An understanding of the genetic and environmental control of flowering time will also assist in the design and management of seed orchards, which are increasingly being used to produce genetically improved *E. globulus* seed for plantations. Synchronous flower opening time among trees in seed orchards maximises open pollination, which is the cheapest way of producing genetically improved seed.

The transition to flowering in plants is the result of the balance of environmental signals and a complex genetic pathway that has been studied extensively in annual plants such as *Arabidopsis*. In perennial trees, which are characterised by a juvenile, non-flowering phase lasting several years, followed by an annual cycling between vegetative and reproductive growth, the genetic control of flowering time is less understood and is potentially more complex. The variation among plant species in their flowering response to environmental signals is important for adaptation to a wide variety of environmental conditions, to ensure that flowering occurs at the most favourable time for reproduction and dispersal. Although much has been learned from *Arabidopsis*, in order to understand the diversity of flowering responses, there is a need to study

flowering pathways in other plant species. Homologues of the *Arabidopsis* flowering genes have been identified in *Eucalyptus* and while it has been shown that these genes are expressed in specific tissue types, to date there have been no studies of how the expression of these genes may be regulated during the processes of flower bud initiation, flower development and flower opening.

This thesis begins by examining the morphological and molecular (microsatellite) variation using samples collected from native trees across the range of the *E. globulus* complex in south-eastern Australia (Chapter 1). The genetic variation in flower bud initiation and opening (flowering time) across the main provenances used in breeding programs (the *globulus* subspecies and some intergrades) is explored in Chapter 2, including the potential role of climate in influencing the year-to-year variation in flower opening time in *E. globulus* trees planted in a common environment. Finally, Chapter 3 identifies genes that may be controlling the variation in seasonal flower bud initiation or opening, by monitoring the seasonal expression of genes that are known to be associated with flowering in the model annual plant *Arabidopsis*. The final Conclusions section briefly summarises the key findings of this thesis.

1 MORPHOLOGICAL AND MOLECULAR VARIATION IN THE *E. GLOBULUS* SPECIES COMPLEX

1.1 Introduction

1.1.1 Population genetics of common and widespread forest trees

To conserve the maximum amount of genetic diversity in a species, thereby maximising its long term adaptive flexibility (Frankel and Soule 1981), the distribution of genetic variation across its geographical and ecological range must be understood. Assessing the genetic variation across the range of widespread species can assist in determining the conservation value of peripheral or disjunct populations. Genetic theory predicts that geographically peripheral or disjunct populations should exhibit lower genetic diversity and higher genetic divergence than central populations, due to the influence of genetic drift and limited gene flow (Hartl and Clark 1997). Isolated populations are also often small, and, as the number of potential mating partners is low, may be subject to inbreeding, which also causes lower diversity. However, a recent review showed that not all species exhibit the expected decline in diversity or increase in differentiation in peripheral populations (Eckert *et al.* 2008), so predictions of genetic structure based on species distribution are not always valid, necessitating analysis on a case-by-case basis.

Common species tend to receive less attention than rare species in conservation biology, despite their importance in many ecosystems (Gaston and Fuller 2008). Many common trees are foundation or keystone species, providing habitat and food for many animals, fungi, micro-organisms and other plants. The emerging field of community and ecosystem genetics has shown that the genetics of foundation plant species can have flow-on effects through the ecosystem (Whitham *et al.* 2008). Studies have shown that the genetics of a range of tree genera such as *Populus*, *Pinus*, *Quercus* and *Salix* (reviewed in Whitham *et al.* 2006, Whitham *et al.* 2009) as well as *Eucalyptus* (Dungey *et al.* 2000, Barbour *et al.* 2008a, Barbour *et al.* 2009a, Barbour *et al.* 2009b), can have extended phenotypic effects which impact on associated communities and even ecosystem processes.

Eucalyptus globulus is a common, widespread tree across south-eastern Australia where it is an important dominant tree of lowland forests and is likely to have a foundation role in ecosystem processes. For example, there is genetically based variation in resistance of *E. globulus* to marsupial herbivores, which is correlated with the levels of plant defensive compounds, sideroxylonals and macrocarpals (O'Reilly-Wapstra *et al.* 2002, 2004; O'Reilly-Wapstra *et al.* 2005). It is, therefore, likely that traits such as defensive chemical concentration have extended phenotype effects. However, until the genes for a range of traits such as these are identified, a conservation strategy based on a population genetics approach, using putatively neutral markers that reflect the overall affinities of individuals and inform us about the presence and extent of barriers to gene flow, will have value in conserving *E. globulus* and its dependent communities.

1.1.2 Distribution and taxonomy of the E. globulus *species complex*

The *Eucalyptus globulus* (bluegum) species complex is naturally distributed in south-eastern mainland Australia and on the island of Tasmania, where it is an important dominant tree of lowland forests (Williams and Potts 1996). It has a largely continuous distribution but there are several small, disjunct populations at the periphery of the species' range (Brooker and Kleinig 1999, Figure 1.1). The complex consists of four taxa variously described as species or subspecies (*bicostata*, *globulus*, *maidenii* and *pseudoglobulus*) (Kirkpatrick 1975; Pryor and Johnson 1971; Jordan *et al.* 1993; Brooker 2000) but most recently as subspecies (Slee *et al.* 2006).

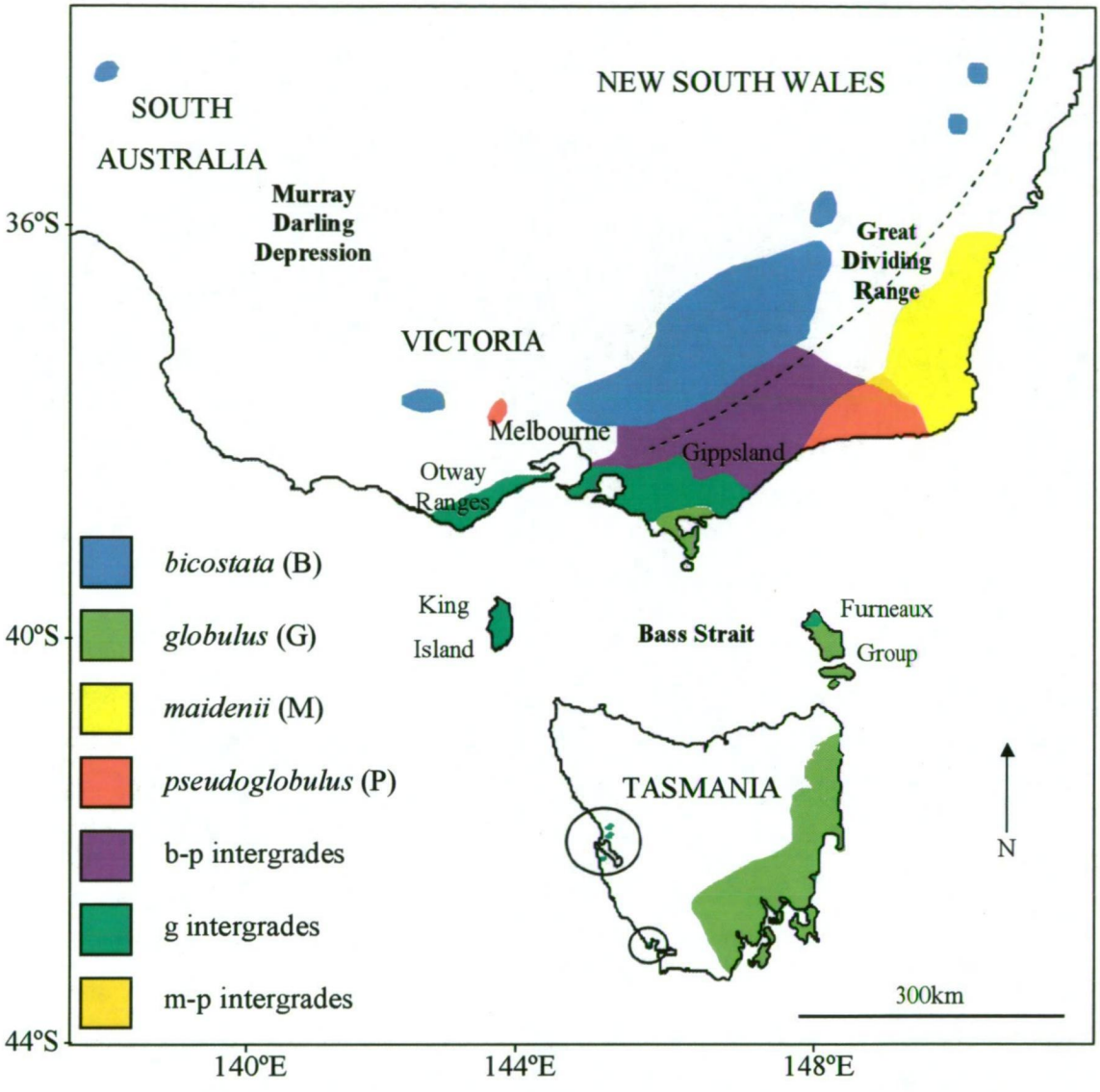


Figure 1.1. The distribution of the *Eucalyptus globulus* subspecies and their intergrades in south-eastern Australia (modified from Jordan *et al.* 1993). b-p intergrades: intermediate between *bicostata* and *pseudoglobulus*; m-p intergrades: intermediate between *maidenii* and *pseudoglobulus*; g intergrades: intermediate between *globulus* and the continuum between *bicostata* and *pseudoglobulus* (Jordan *et al.* 1993). Within the two circles are small g intergrade populations. The three major geographic barriers are shown in bold: the Murray Darling Depression (lowland plains with a warm, semi-arid climate), the Great Dividing Range (indicated by a dashed line) and the Bass Strait (channel that separates the island of Tasmania from the rest of Australia).

The four taxa are differentiated mainly by capsule morphology traits. Subspecies *maidenii* has the smallest fruit of the complex with up to seven fruit per umbel, and occurs on the south coast of New South Wales (NSW). Subspecies *bicostata* and *pseudoglobulus* are both three-fruited, with *pseudoglobulus* having smaller fruit, longer pedicels and fewer ribs on the capsules than *bicostata*. Core *pseudoglobulus* mostly occurs in East Gippsland, while *bicostata* occurs inland of the Great Dividing Range in Victoria (Vic.) and NSW and a disjunct population in South Australia (SA). Subspecies *globulus* has large, solitary capsules, and occurs through eastern Tasmania (Tas.), most of the Furneaux group of islands and the coast of South Gippsland. The cores of these four taxa are believed to be morphologically as well as geographically distinct, but are linked by intergrade populations that are intermediate in morphology (Kirkpatrick 1975). These intergrade populations result in a complete continuum in capsule morphology across the geographic range of *E. globulus* which means that trees are often misclassified in the field. This has implications for seed collections for plantations and breeding programs. In addition, the evolutionary origins of the intergrade populations are unknown and may be the result of: a) primary differentiation (divergence within a continuous series of populations along a selective gradient), or b) secondary intergradation (hybridisation and introgression between previously isolated gene pools). These two evolutionary scenarios are difficult to distinguish based on current patterns of variation (Endler 1977). However, a recent zone of secondary intergradation would be expected to display increased diversity due to gene flow from differentiated populations (Rieseberg and Wendel 1993). For example, molecular data show that the genetic diversity within a *Populus fremontii* - *P. angustifolia* hybrid zone is much higher than in the pure species zones (Whitham *et al.* 1999). In addition, a recent hybrid zone would be expected to have correlated morphological and molecular affinities, as was the case in the *Quercus crassifolia* - *Q. crassipes* hybrid zone in Mexico (Tovar-Sanchez and Oyama 2004).

1.1.3 Geographic barriers to dispersal in the E. globulus *species complex*

Gene flow in *E. globulus*, like other forest trees, is more a function of pollen than seed dispersal. *Eucalyptus globulus* seed is relatively heavy and poorly adapted for dispersal (Boland *et al.* 1980) suggesting that it is likely that most seed is deposited within a

radius of twice the canopy height (Potts and Wiltshire 1997). Molecular studies have confirmed that seed dispersal in *E. globulus* is probably limited (Skabo *et al.* 1998, Jones *et al.* 2007). On the other hand, pollen is dispersed by active pollinators such as birds and insects (Hingston *et al.* 2004a; Hingston *et al.* 2004b) and while most pollen is also dispersed within close proximity to the source, the frequency of long distance pollen dispersal events is higher than that of seed (Potts and Wiltshire 1997). As chloroplast DNA is maternally inherited in eucalypts (Byrne *et al.* 1993), comparisons of differentiation statistics for nuclear markers *versus* chloroplast markers in the same species can give an indirect measure of the relative importance of pollen *versus* seed flow. In nearly all eucalypt species that have been analysed in this way, pollen tends to be much more widely dispersed than seed (reviewed in Byrne 2008).

As seed dispersal in *E. globulus* is limited, geographic barriers across its range will provide significant barriers to dispersal. The major geographic barriers in the *E. globulus* range are the Bass Strait, the Great Dividing Range and the Murray Darling Depression (see Figure 1.1).

The Bass Strait is a 250 km wide channel that separates the island of Tasmania from the rest of Australia. It is relatively shallow (mostly less than 100 m deep) and therefore changes in sea level have caused Tasmania to be repeatedly separated and rejoined to mainland Australia by the Bassian Plain. This land bridge linked north-eastern Tasmania to South Gippsland through the Furneaux Group, and, at times, linked north-western Tasmania to the Melbourne region through King Island (Jackson 1999).

The Great Dividing Range is a mountain range that stretches 3 500 km down the east coast of Australia, separating the coastal strip from the inland plains. The Australian Alps, the highest part of the Great Dividing Range with peaks of over 2 000 m, bisects the *E. globulus* distribution in north-eastern Victoria / southern NSW (Figure 1.1). This altitude is outside the tolerance of the *E. globulus* species complex, which currently occurs to a maximum altitude of 1 100 m (Kirkpatrick 1973) but would have retreated even further downslope during glacial periods: it has been suggested that most of the central highlands of Victoria was treeless during the last glacial maximum (McKenzie 1997) and the Tasmanian treeline was probably close to present sea level (Kirkpatrick and Fowler 1998). While significant portions of Tasmania were covered by ice in

previous glacial periods (1 100 km^2 in the last glacial maximum), there were only small areas of glaciations in the southern NSW portion of the Australian Alps (Barrows *et al.* 2001), and there is little evidence of glaciation in the Victorian Highlands (McKinnon *et al.* 2004a).

The Murray Darling Depression which extends through north-western Victoria, southwestern NSW and eastern South Australia, is a lowland plain with a warm, semi-arid climate dominated by grassland and mallee (multi-stemmed) eucalypts. The current aridity of this region makes it unsuitable habitat for the *E. globulus* complex, and the area would have been even more arid during the last glacial maximum. Also, from the late Eocene to the mid Miocene there was a period of marine incursion (Nelson 1981) including, at times, into the Murray Darling Depression, which would have also prevented migration of the *E. globulus* complex between Victoria and South Australia. However, evidence from Lake Eyre suggests there may have been wetter periods between 50 000 and 35 000 years ago (Magee and Miller 1998) and the possibility that the Murray Darling Basin supported more or less continuous populations of the *E. globulus* species complex during this or previous wetter periods cannot be dismissed.

1.1.4 Variation in the E. globulus *species complex*

Variation across all or part of the species complex has been investigated in several taxonomic (Jordan *et al.* 1993; Kirkpatrick 1975), quantitative (Jordan *et al.* 1994; Potts and Jordan 1994, Dutkowski and Potts 1999; MacDonald *et al.* 1997, Jordan *et al.* 1999, Jordan *et al.* 2000, Muneri and Raymond 2000; O'Reilly-Wapstra *et al.* 2002, Whittock *et al.* 2003) and molecular (Freeman *et al.* 2001; Jones *et al.* 2002; Nesbitt *et al.* 1995; Steane *et al.* 2006) genetic studies (reviewed in Potts *et al.* 2004) to elucidate evolutionary relationships between the taxa and populations, and to assist in the prediction of breeding values.

Kirkpatrick (1975) measured a range of morphological traits in samples collected across the natural range of *E. globulus* and showed that there was morphological variation in the adult phenotype across the species range, with four differentiated taxa and extensive zones of clinal intergradation between the taxa. He concluded that *E. globulus* is a single species with four subspecies, due to the continuous nature of the morphological

variation, and identified *pseudoglobulus* as the central group of the complex with the other three taxa intergrading into it. It was suggested that the intergrade populations arose mostly through primary differentiation rather than secondary intergradation, as, for secondary intergradation to occur, isolation and recontact would have had to occur in three directions (i.e. between *pseudoglobulus* and the other three taxa). However he did not discount the possibility of a combination of both primary and secondary intergradation occurring in the complex.

Jordan *et al.* (1993) also measured a range of morphological traits, combining the Kirkpatrick (1975) collections with CSIRO Australian Tree Seed Centre collections of ssp. *globulus* and its intergrades (Gardiner and Crawford 1987, 1988). This analysis confirmed the presence of four differentiated groups in *E. globulus*, corresponding to the four subspecies, and also showed a complete continuum between *bicostata* and *pseudoglobulus*, and between *maidenii* and *pseudoglobulus*, in terms of capsule morphology. However, the authors suggested that *globulus* intergraded with the *pseudoglobulus-bicostata* continuum, rather than into *pseudoglobulus*. Jordan *et al.* (1993) concluded that a combination of primary differentiation and secondary intergradations had acted to shape the observed patterns of morphological diversity in the species complex, but conceded that selection acting on morphological traits could cause a misrepresentation of true genetic relationships in the species complex.

Dutkowski and Potts (1999) assessed 35 quantitative traits in five common environment field trials of *globulus* and its intergrades. This analysis revealed high levels of spatially structured genetic variation, which was summarised by grouping localities into 13 races and 20 subraces. This classification proved useful for predicting breeding values as it included a number of economically important traits. However, classifications using quantitative genetics have their limits since these quantitative traits are under the influence of selection (Steane *et al.* 2006) and may not depict the true genetic relationships among populations of *globulus* and its intergrades.

Molecular variation in the species complex has been analysed using Random Amplification of Polymorphic DNA (RAPD) markers (Nesbitt *et al.* 1995), chloroplast DNA (cpDNA) (Freeman et al. 2001) and microsatellite markers (Jones et al. 2002, Steane et al. 2006). As cpDNA is maternally inherited in eucalypts (Byrne *et al.* 1993),

including *E. globulus* (McKinnon *et al.* 2001b), patterns of cpDNA variation reflect patterns of seed dispersal and can be used to identify migratory routes. Based on the distribution of cpDNA haplotypes in *globulus*, Freeman *et al.* (2001) suggested a western migration route between Victoria (Otway Ranges) and Western Tasmania via King Island, and a barrier to seed dispersal between the Furneaux Group and Tasmania. Studies of nuclear genetic variation in the *E. globulus* complex also revealed high levels of spatially structured genetic diversity (Jones *et al.* 2002; McKinnon *et al.* 2005; Nesbitt *et al.* 1995; Steane *et al.* 2006), however very few, or no, samples of *pseudoglobulus*, *bicostata* and *maidenii* were included in these studies.

1.1.5 Using molecular markers to assess taxonomic affinities in species complexes

The most widely accepted species concept defines species as reproductively isolated units (Dobzhansky 1970). However, for many organisms, morphological and ecological differences among recognised species are not always accompanied by reproductive isolation (Stebbins 1989) and morphological boundaries between species can be blurred by hybridisation and introgression, resulting in species complexes that confound taxonomic classification. This is the case in herbaceous species such as sunflowers and orchids (Rieseberg *et al.* 1991, Dressler and Dodson 1960) as well as forest tree genera such as *Quercus*, *Pinus*, *Populus* and *Eucalyptus* (Johnson 1976, Manos and Fairbrothers 1987, Griffin *et al.* 1988). In *Eucalyptus*, incomplete morphological resolution of taxa and the recognition of many subspecies is thought to be due to recent and ongoing speciation (Byrne 2008).

The use of morphological traits to define taxa in a species complex can be problematic, as homology (similarity due to common ancestry) and homoplasy (similarity due to convergence) cannot be distinguished. In the last few decades, isozymes and DNA markers have been used to assess genetic differentiation and the evolutionary history of species complexes, and such studies have often identified low levels of genetic divergence among taxa, implying recent divergence and/or hybridisation and introgression. Particularly well-studied taxa include sunflowers (e.g. Beckstrom-Sternberg *et al.* 1991, Rieseberg *et al.* 1991, Yatabe *et al.* 2007) and oaks (e.g. Hokanson *et al.* 1993, Hess and Stoynoff 1998, Gomory *et al.* 2001, Kashani and Dodd

2002, Aldrich *et al.* 2003). Molecular markers have also been used to elucidate phylogenetic relationships among taxa in many forest tree species complexes such as the *Abies magnifica-A. procera* (Oline 2008), *Melaleuca quinquenervia* (Cook *et al.* 2008), *Astronium* (Caetano *et al.* 2008), *Fagus sylvatica* (Gomory *et al.* 2007) and *Acacia acuminata* (Byrne *et al.* 2002) species complexes. Recent advances in Bayesian assignment procedures have proven particularly useful in defining taxonomic units in species complexes (e.g. Caetano *et al.* 2008, Gomory *et al.* 2007). Duminil *et al.* (2006) even argued that the best approach when studying related taxa is to use only Bayesian procedures for species delimitation, without reference to morphological data.

In *Eucalyptus*, isozyme and DNA methods have also been used widely, usually in combination with morphological analyses, to differentiate between taxa and determine taxonomic status, or to detect hybridisation in species complexes. Examples include *E. saligna-E. botryoides* (Passioura and Ash 1993), *E. argutifolia-E. obtusiflora* (Kennington and James 1998), *E. risdonii – E. tenuiramis* (Turner 2000), *E. vernicosa-E. subcrenulata-E. johnstonii* (McGowen *et al.* 2001), the *E. loxophleba* complex (Hines and Byrne 2001), *E. populnea - E. brownii* (Holman *et al.* 2003), *E. kochii – E. horistes* (Byrne 1999) and the *E. angustissima* complex (Elliott and Byrne 2004). Similar to studies of the sunflower and oak species complexes, low levels of genetic divergence are commonly reported among closely related species of eucalypts. In some cases (e.g. Passioura and Ash 1993, Byrne 1999) a reduction in taxonomic status from species to subspecies was suggested, while other studies have raised subspecies to species rank (e.g. Elliott and Byrne 2004).

1.1.6 Gene pool management of E. globulus

Eucalyptus globulus is the most widely planted hardwood species in temperate parts of the world (Eldridge *et al.* 1993). Australia is, therefore, a custodian of an internationally important genetic resource and has a responsibility to safeguard the native gene pool by developing sustainable management practices. Intelligent strategies are especially important, in the current era of rapid climate change and changing breeding objectives. A gene pool management strategy for this species will maintain genetic diversity for

seed collections for breeding programs and plantations, and identify and protect populations of conservation significance.

As well as being planted worldwide, *E. globulus* is planted in parts of south-eastern Australia including areas within its own natural range (Barbour *et al.* 2008b). As some of the taxa of the species complex are easily crossed artificially (Potts *et al.* 1992) and the main provenances used in plantations are usually from particular localities within the natural range (due to their superior performance in field trials), there is concern that there is potential for gene flow from non-local plantings into the native gene pool (Potts *et al.* 2003). The impact that such gene flow will have on native populations will depend on the genetic relationship between the plantation provenance and the local population, since the more different they are, the more exotic the genetic contribution from plantations will be. Similarly, the impact of seed transfer for reforestation programs will depend on the origin and genetic diversity of the transplanted material (Lefevre 2004).

1.1.7 Aims of this study

The objectives of this study were to use nuclear DNA markers (microsatellites) coupled with morphological measurements to characterise the broad scale patterns of genetic variation across the *E. globulus* species complex, to provide insights into some of the evolutionary processes that have shaped the complex patterns of genetic variation in the *E. globulus* gene pool, to evaluate the conservation status of small isolated populations, and to provide baseline data for a gene pool management strategy for the species complex.

1.2 Materials and Methods

1.2.1 Plant material

Leaf tissue for DNA extraction, a herbarium specimen (including buds and capsules) and open pollinated seed (where available) were sampled from 565 adult trees representing 50 natural populations, 17 regions and three taxa of the *E. globulus* species complex, including their intergrades (Table 1.1, Figure 1.2). Each tree was tagged and photographed. Trees sampled were at least 100 m apart, to avoid sampling closely related trees (see Skabo *et al.* 1998, Jones *et al.* 2007). Localities of *globulus* and its intergrades were grouped into races (as defined by Dutkowski and Potts 1999), hereafter referred to as regions; the other taxa and their intergrades were grouped into geographic regions based on overall morphology (Table 1.2). Populations were classified into a pure taxon or intergrade type following the Brooker and Kleinig (1999) field guide and/or the collections of Kirkpatrick (1975) and Jordan *et al.* (1993) (Table 1.2). Populations in the Buchan region (region 15), classified as pure *pseudoglobulus* or intergrade *bicostata-pseudoglobulus* by Jordan *et al.* (1993) appeared to also have affinities to *maidenii* when collected in this study, so this region was classified as "*bicostata-maidenii-pseudoglobulus*" intergrade (b-m-p) (Table 1.2).

Table 1.1 Samples of the *E. globulus* complex used in this study, and the number (*n*) of core taxa, morphological groups of intergrades, regions and localities represented. Final number (*n*) of individuals is the number of individuals remaining after excluding those that were missing data from three or more loci. For taxon codes see Table 1.2.

Study	Taxa sampled	*n* (taxa)	*n* (intergrade types)	*n* (regions)	*n* (localities)	*n* (individuals)
This study	M, B, P, m-p, b-m-p, b-p	3	3	17	49	556
Jones *et al.* 2002	G, B, P, b-p, b-m-p, g	3	3	8	18	117
Steane *et al.* 2006	G, g	1	1	14	38	443
Foster *et al.* 2007	G	1	0	3	5	110
Jones *et al.* 2007	G	1	0	1	1	9
Combined data	G, B, P, M, m-p, b-m-p, b-p, g	4	4	33	98	1235

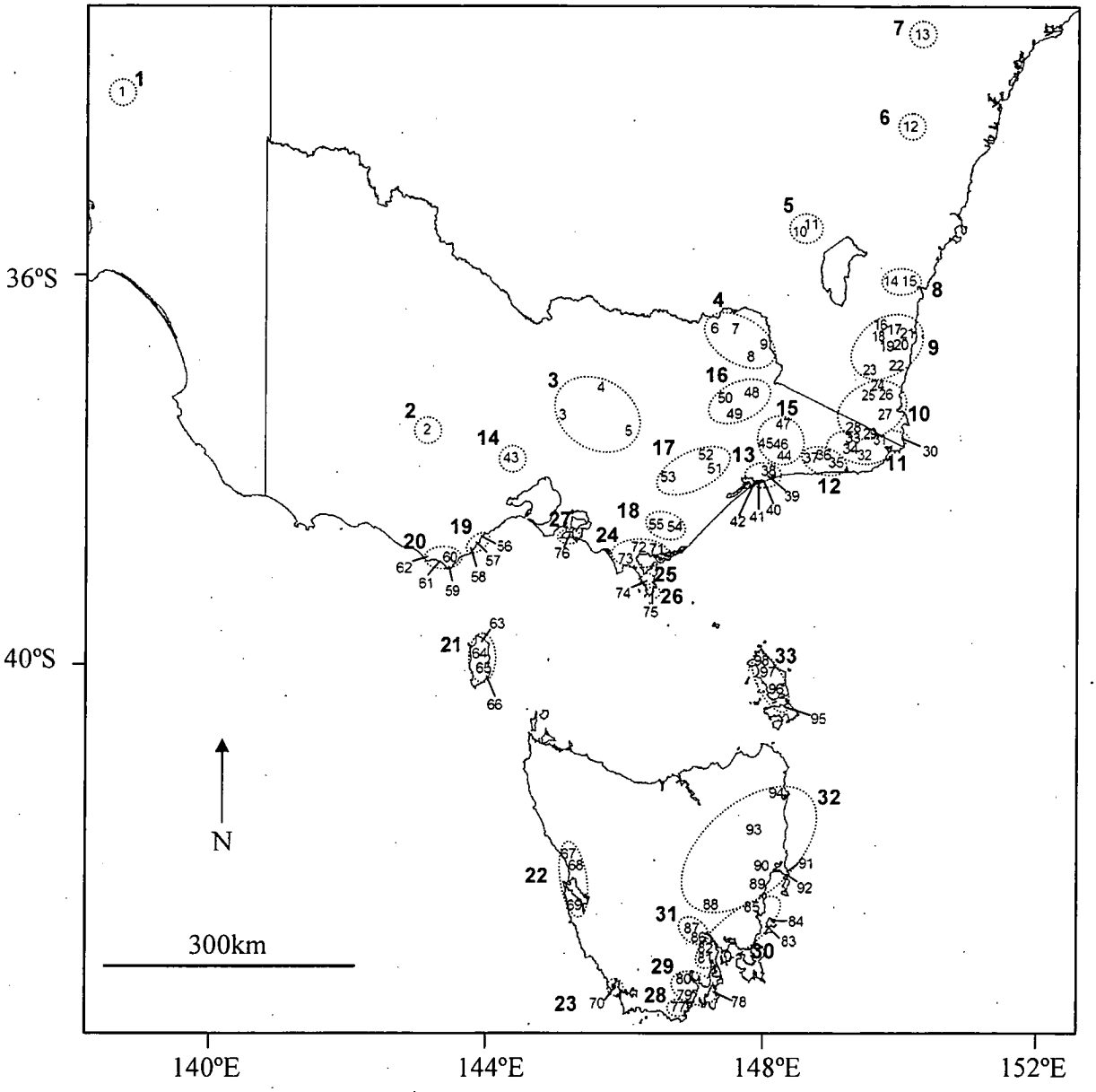


Figure 1.2. Localities and regions of the *E. globulus* complex used in this study. Locality positions are given by numbers; regions are circled and numbered in bold. Localities and regions are named and detailed in Table 1.2.

Table 1.2. Sample sizes (*n*) of *E. globulus* localities in this study, after excluding individuals that were missing data from three or more loci.

	Region		Locality					
Taxon (Code)	Code	Name	Code	Name	Latitude	Longitude	Altitude	*n*
bicostata (B)								
	1	Mt Bryan	1	Mt Bryan	-33.4398	138.9585	821	25
	2	Mt Cole	2	Mt Cole	-37.2970	143.2439	826	19
	3	Hume	3	Tallarook	-37.1525	145.1195	479	9
		Hume	4	Euroa	-36.8239	145.6923	600	12
		Hume	5	Eildon	-37.3081	146.0895	850	9
	4	NE Victoria	6	Mt Granya	-36.1349	147.2714	867	8
		NE Victoria	7	Shelley	-36.1816	147.5476	754	8
		NE Victoria	8	Nariel	-36.4840	147.7981	1029	8
		NE Victoria	9	Mt Unicorn	-36.3400	147.9670	843	8
	5	Canberra	10	Bungongo	-35.0450	148.4842	791	13
		Canberra	11	Burrinjuck	-34.9654	148.6298	505	7
	6	Jenolan	12	Jenolan	-33.8195	150.0267	771	13
	7	Wollemi	13	Nullo Mountain	-32.7354	150.2234	1102	20
bicostata total								159
maidenii (M)								
	8	Araluen	14	Araluen	-35.6062	149.7947	471	16
		Araluen	15	Currowan-Monga	-35.6066	150.0526	383	23
	9	Wadbilliga	16	Wadbilliga North	-36.1255	149.6331	353	14
		Wadbilliga	17	Belowra Rd	-36.1592	149.8380	492	10
		Wadbilliga	18	Wadbilliga South	-36.2508	149.6392	329	20
		Wadbilliga	19	Murrabrine	-36.3568	149.7441	428	20
		Wadbilliga	20	Kooraban NP	-36.3422	149.8765	247	6
		Wadbilliga	21	Mt Dromedary (NSW)	-36.2869	150.0395	220	13
		Wadbilliga	22	Mumbulla SF	-36.5790	149.8576	185	20
		Wadbilliga	23	Brown Mountain	-36.6125	149.4748	436	12
	10	SE Forests	24	Tantawangalow	-36.8021	149.5771	299	8
		SE Forests	25	Big Jack - Rocky Hall	-36.8930	149.4556	357	15

Taxon (Code)	Region Code	Region Name	Locality Code	Locality Name	Latitude	Longitude	Altitude	*n*
		SE Forests	26	Yurammie	-36.9000	149.7194	480	16
		SE Forests	27	Mt Imlay	-37.1216	149.6904	316	8
		SE Forests	28	north Cann Valley Highway	-37.2680	149.2293	298	7
		SE Forests	29	Wroxham	-37.3442	149.4739	297	8
maidenii total								216
maidenii - pseudoglobulus intergrade (m-p)								
	11	Alfred-Nadgee	30	Nadgee	-37.4376	149.9550	14	18
		Alfred-Nadgee	31	Maramingo Creek	-37.4163	149.6241	173	10
		Alfred-Nadgee	32	Alfred NP	-37.5672	149.3400	331	9
		Alfred-Nadgee	33	mid Cann Valley Highway	-37.4320	149.1968	151	6
		Alfred-Nadgee	34	south Cann Valley Highway	-37.5008	149.1798	132	12
maidenii - pseudoglobulus intergrade total								55
pseudoglobulus (P)								
	12	Mt Cann	35	Mt Cann	-37.6539	148.9650	311	10
		Mt Cann	36	Wiebens Hill	-37.6074	148.7799	282	8
		Mt Cann	37	Murrungowar	-37.6242	148.6939	252	8
	13	Lakes Entrance	38	Ostlers Rd	-37.8003	148.0394	43	9
		Lakes Entrance	39	Lake Tyers	-37.8552	148.0410	25	14
		Lakes Entrance	40	Lakes Entrance	-37.8798	147.9744	41	11
		Lakes Entrance	41	Metung-Lakes Entrance	-37.8825	147.8930	43	6
		Lakes Entrance	42	Metung	-37.8795	147.8573	24	8
	14	Lerderderg Gorge	43	Lerderderg Gorge	-37.6112	144.4165	153	20
pseudoglobulus total								94
bicostata-maidenii-pseudoglobulus intergrade (b-m-p)								
	15	Buchan	44	Stoney Creek	-37.5705	148.2556	241	10
		Buchan	45	Cutts Creek Rd	-37.4714	148.0115	438	8
		Buchan	46	Buchan	-37.4518	148.1899	339	11
		Buchan	47	Gelantipy	-37.2652	148.2491	746	8
bicostata-maidenii-pseudoglobulus intergrade total								37

Taxon (Code)	Region Code	Region Name	Locality Code	Locality Name	Latitude	Longitude	Altitude	*n*
bicostata-pseudoglobulus intergrade (b-p)								
	16	Omeo	48	Beloka Rd	-36.8930	147.8160	953	7
		Omeo	49	Omeo	-37.1154	147.5618	763	8
		Omeo	50	Trapyard Gap	-36.9376	147.4343	1096	7
	17	Mitchell River	51	Cobbannah	-37.6934	147.2891	400	10
		Mitchell River	52	Peel Gap	-37.5928	147.1833	514	10
		Mitchell River	53	Heyfield[1]	-37.8241	146.6681	290	4
bicostata-pseudoglobulus intergrade total								46
globulus intergrade Victoria (g)								
	18	Strzelecki Ranges	54	Bowden-Carrajung[2]	-38.6318	146.6988	365	2
		Strzelecki Ranges	55	Jeeralang	-38.3509	146.4866	140	23
	19	Eastern Otways	56	Lorne	-38.5327	143.9700		14
		Eastern Otways	57	Jamieson Creek	-38.6012	143.9116		8
		Eastern Otways	58	Cape Patton	-38.6723	143.8527		7
	20	Western Otways	59	Parker Spur	-38.8118	143.5558		8
		Western Otways	60	Cannan Spur	-38.7629	143.5377		8
		Western Otways	61	Otway State Forest	-38.7758	143.4196		8
		Western Otways	62	Lavers Hill	-38.7448	143.2541		8
globulus intergrade Tasmania (g)								
	21	King Island	63	North King Island	-39.6528	144.0417		6
		King Island	64	Central North King Island	-39.7589	144.0017		5
		King Island	65	Central King Island	-39.9429	144.0278		22
		King Island	66	South King Island	-40.0297	144.0633		7
	22	Western Tasmania	67	Little Henty River	-41.9415	145.1976	58	8
		Western Tasmania	68	Badgers Creek	-42.0016	145.2751	115	8
		Western Tasmania	69	Macquarie Harbour	-42.3371	145.3310	3	16
	23	Port Davey	70	Port Davey	-43.2977	145.9226	19	35
globulus intergrade total								193
globulus Victoria (G)								
	24	South Gippsland	71	Alberton West-Welshpool-Hedley	-38.6344	146.4853	20	14
		South Gippsland	72	Toora	-38.6722	146.2803	20	8
		South Gippsland	73	Fish Creek	-38.7178	146.1017		6

Taxon (Code)	Region		Locality					
	Code	Name	Code	Name	Latitude	Longitude	Altitude	*n*
	25	Tidal River	74	Tidal River	-39.0212	146.3353		31
	26	Wilson's Promontory	75	Wilson's Promontory Lighthouse	-39.1179	146.4240	75.8	31
	27	Phillip Island	76	Phillip Island	-38.4814	145.2576	32	24
globulus Tasmania (G)								
	28	Recherche Bay	77	Recherche Bay	-43.5274	146.8963		27
	29	Southern Tasmania	78	South Bruny Island	-43.3498	147.3208	5	8
		Southern Tasmania	79	Dover	-43.3052	147.0730		7
		Southern Tasmania	80	Geeveston	-43.1918	146.8960	100	10
	30	SE Tasmania	81	Tinderbox	-43.0332	147.3196	300	9
		SE Tasmania	82	Hobart[2]				1
		SE Tasmania	83	Maria Island South	-42.7171	148.0721	82	30
		SE Tasmania	84	Maria Island North	-42.5850	148.0668	62	6
		SE Tasmania	85	Triabunna	-42.4620	147.8960	265	8
	31	Dromedary Tas	86	Mt Dromedary (Tas)	-42.7209	147.1324	389	19
		Dromedary Tas	87	Platform Peak	-42.6766	147.0302	539	17
	32	NE Tasmania	88	Jericho	-42.4254	147.2649		6
		NE Tasmania	89	Mayfield South	-42.2125	148.0313	50	6
		NE Tasmania	90	Mayfield North	-42.0177	147.9719	296	8
		NE Tasmania	91	Cape Tourville Dwarf	-42.1230	148.3428	122	20
		NE Tasmania	92	Cape Tourville Tall	-42.1210	148.3375	99	38
		NE Tasmania	93	Pepper Hill	-41.6380	147.8465	602	8
		NE Tasmania	94	Humbug Hill	-41.2690	148.3103	139	9
	33	Furneaux	95	Cape Barren Island	-40.3413	148.2595		27
		Furneaux	96	Flinders Island south[2]				3
		Furneaux	97	Flinders Island central	-39.9530	148.0019	35	15
		Furneaux	98	Flinders Island north[2]				2
globulus total								398
E. globulus Grand Total								1198

[1] Heyfield was excluded from the regional and STRUCTURE analysis but not the locality level analysis as it was geographically distant from other populations of similar morphology.

[2] Due to small sample size, these four localities were excluded from analyses at the locality level, but were included in the STRUCTURE and regional level analyses.

1.2.2 Molecular methods

1.2.2.1 DNA extraction and PCR

Total genomic DNA was extracted from fresh or frozen leaf tissue using a CTAB method (Doyle and Doyle 1990) with several modifications (McKinnon *et al.* 2004b). DNA quality and quantity were assessed by gel electrophoresis and comparison with a lambda *Hin*dIII molecular weight standard.

Initially, 11 nuclear microsatellite loci were tested; however, due to inconsistencies with scoring, EMBRA10 and EMCRC3 were excluded. Of the nine loci selected, six were designed by Steane *et al.* (2001), two by Brondani *et al.* (1998) and one by Brondani *et al.* (2002), and the primer sequences are available therein. These loci have been mapped and there is no evidence for linkage between the loci (J. Freeman pers. comm.). Forward primers were synthesised with a WellRED fluorescent label (Proligo) to enable detection on a CEQ fragment analysis system (Beckman Coulter™). PCR conditions followed Brondani *et al.* (1998) for the three EMBRA loci, and Jones *et al.* (2002) for the EMCRC loci except that 0.28 μM of each primer was used in a total reaction volume of 12.5 μL. PCRs were performed separately and PCR products co-loaded in two sets (Set 1: EMCRC5, EMCRC6, EMCRC7, EMBRA19; Set 2: EMCRC2, EMCRC10, EMCRC11, EMBRA11, EMBRA30) and sized by comparison with the CEQ™ DNA Size Standard-400 using the Beckman Coulter™ Fragment Analysis software for allele binning.

A set of ten DNA samples from previous studies were included as controls to allow data sets from Jones *et al.* (2002), Steane *et al.* (2006), Jones *et al.* (2007) and Foster *et al.* (2007) to be combined with results from this study, with allele sizes adjusted where necessary. These ten samples were chosen as they contained the minimum and maximum allele sizes and the most common alleles at each locus. Steane *et al.* (2006) used only EMCRC loci, however, their samples have since been genotyped using the four EMBRA loci used in this study (Steane, unpublished data). Within this study, 100 PCRs were duplicated (representing all nine loci). Only samples with results for six or more loci were retained, resulting in a final data set of 1198 individuals representing 98 localities, 33 regions and all four taxa of the *E. globulus* species complex, including their intergrades, across the entire natural range of the species complex (Figure 1.1, Table 1.2).

1.2.2.2 Molecular data analysis

1.2.2.2.1 Descriptive analysis

All statistics calculated at the locality or regional level were based on the 94 localities or the 33 regions listed in Figure 1.2 and Table 1.2. GDA 1.1 (Lewis and Zaykin 2001) was used to calculate the following genetic diversity parameters, averaged for each locus, locality and region: number of alleles (A), observed and expected heterozygosity (H_O and H_E, respectively) and Wright's Fixation Index (F), linkage disequilibrium among loci was also tested. Null allele frequencies in each region were estimated using the inbreeding population method in INEst (Chybicki and Burczyk 2009). Allelic richness (El Mousadik and Petit 1996) was calculated at the locality and region levels, using FSTAT (Goudet 2002). The proportion of loci with an allele greater than 0.5 in frequency was also calculated. Contours of allelic richness and expected heterozygosity per locality and region were plotted on a map using 3DField 2.9.6 (Galouchko 2007).

A pairwise matrix of Nei's (1972) Genetic Distance among regions, calculated using GenAlEx (Peakall and Smouse 2006) was used to construct a Neighbour Joining (Saitou and Nei 1987) radial tree with proportional branch lengths, using T-REX 4.0 (Makarenkov 2001).

F-statistics (F_{IS}, F_{IT}, F_{ST}, Weir and Cockerham 1984) were calculated for each locus at the locality and regional levels, using FSTAT (Goudet 2002). Jost's (2008) estimated D (D_{est}) for each locus at the locality and regional levels was calculated using SMOGD (Crawford 2009). D_{est} is a recently developed measure of differentiation that replaces G_{ST} and its relatives (e.g. F_{ST}). D_{est}, unlike G_{ST}, accounts for differences in diversity among populations and types of loci (Jost 2008) and therefore is a better statistic for comparison among studies, or among taxa within a study. F_{ST} and D_{est} were also calculated within each morphological core taxon, based on the core taxa as defined in Table 1.2. The hierarchical partitioning of genetic variation among localities, regions and taxa was illustrated with F-statistics and AMOVA (Analysis of Molecular Variance, Excoffier *et al.* 1992; Michalakis and Excoffier 1996) using GenAlEx (Peakall and Smouse 2006). Two separate three-level analyses were used to enable the analysis of a four-level hierarchy; the first quantified the proportion of genetic variation among taxa,

among regions within subspecies, and among individuals within regions; the second analysis quantified the variation among regions, among localities within regions, and among individuals within localities. The results of the two analyses were combined as in Jones *et al.* (2006), by partitioning the variation among individuals within regions (from the first analysis) into the variance among individuals within localities, and the variance among localities within regions, from the second analysis.

To determine the relationships among the four taxa, only individuals belonging to "core" subspecies localities were included in the AMOVA analysis. The cores were based on morphology, geography and prior analyses with STRUCTURE (see 1.2.2.2.2) (i.e. individuals from geographically disjunct localities, morphological intergrade localities, and intermediate localities from STRUCTURE analyses were excluded from the analysis). This resulted in a sample size of 567 individuals from 45 localities, 12 regions and four subspecies/taxa (Figure 1.3, Appendix 1). Significance levels were determined after 999 random permutations of individuals in the data set. The same subset of "core" individuals was used to calculate pairwise differentiation between taxa. Pairwise F_{ST} and Nei's (1978) genetic distance were calculated among taxa using the same subset of "core" individuals, but including Victorian core *globulus* (24-South Gippsland and 27-Phillip Island) as well as Tasmanian core *globulus*.

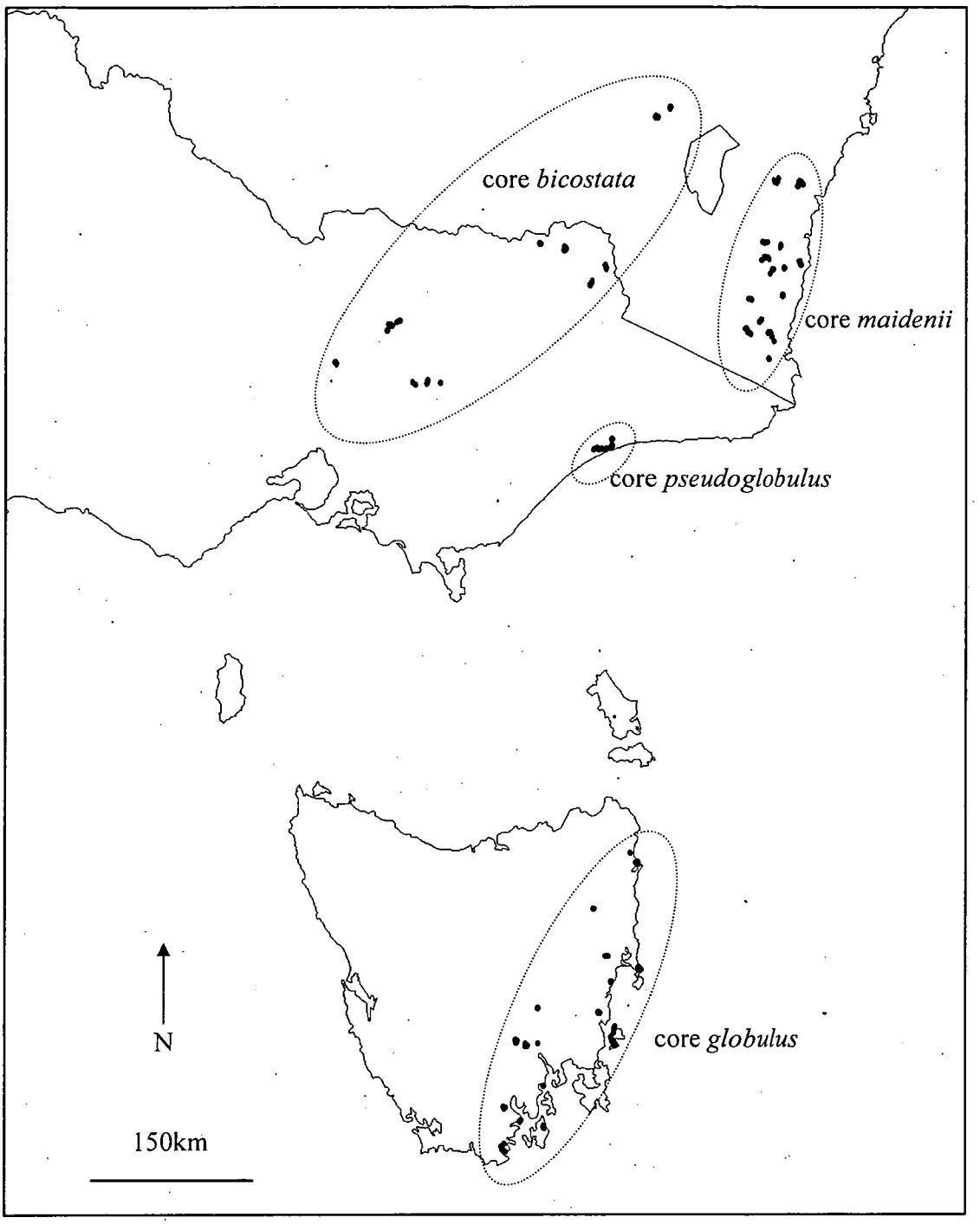


Figure 1.3. Distribution of the subset of "core" samples used for the AMOVA of the *E. globulus* species complex (see Appendix 1 for names and sample sizes of localities and regions that were included).

1.2.2.2.2 Bayesian analysis

The number of groups of genetically similar individuals (K) in the *E. globulus* complex, and the affinities of individuals to these groups (Q), were determined using STRUCTURE 2.2.3 (Pritchard *et al.* 2000) and the ΔK method described in Evanno *et al.* (2005). Assuming no prior population groupings and using the admixture model, the estimated K was determined by comparing the estimated log probability of data at different values of K (from K=1 to K=25), using 100 000 MCMC repetitions following a burnin of 120 000 repetitions (at which point stationarity had been reached). Five independent runs for each value of K were performed, and results were interpreted using the three runs with the highest likelihood.

In a separate set of analyses, designed to examine the affinities of the intergrade populations of mainland Australia, all Tasmanian populations and geographical outliers on mainland Australia were excluded from the full data set, resulting in a data set of 749 individuals representing 21 regions (Appendix 2). STRUCTURE analyses were performed using the same ancestry model as above: an admixture model with no *a priori* population groupings (testing K=1 to K=25), using 100 000 MCMC repetitions but with a burnin of 100 000 repetitions (at which point stationarity had been reached).

Independent runs of STRUCTURE at a given K can result in different solutions, simply due to label-switching across different runs, or due to true multimodality (distinct solutions across different runs). Jakobsson and Rosenberg (2007), in the computer program CLUMPP, developed three algorithms to overcome these problems, which combine independent runs into a single solution. The *Fullsearch* algorithm in CLUMPP was used to derive a single output from the three independent runs at each K from K=2 to K=7. DISTRUCT (Rosenberg 2004) was used to display the probability of membership (Q) of each individual into each of the inferred clusters. The average proportion of membership of each of the 33 regions into the K clusters, from the analysis of all sampled individuals of the *E. globulus* complex, was also plotted on a separate map for each level of K.

A third STRUCTURE analysis was undertaken to directly compare molecular and morphological (see 1.2.3) affinities of individuals. This analysis used a model with *a*

priori population information for populations of "core" *globulus*, *bicostata*, *pseudoglobulus* and *maidenii* (Appendix 3), with an admixture model used for intergrade and geographically outlying populations, to assign these individuals into the clusters defined as core *globulus*, *bicostata*, *pseudoglobulus* and *maidenii*. The original "core" individuals were also reassigned to the four clusters using an admixture model. These core individuals were therefore duplicated in the analysis. Ten independent runs at K=4 were performed using 50 000 MCMC repetitions and a burnin of 50 000 repetitions, and the individual Q values from the run with the highest likelihood were used.

1.2.3 Morphological methods

1.2.3.1 Morphological measurements

For 446 of the trees sampled in this study, 14 capsule, inflorescence and leaf variables were measured (Figure 1.4, Table 1.3) following the descriptions of Kirkpatrick (1975), Jordan *et al.* (1993) and Jones *et al.* (2002). Five average-sized umbels with all parts intact were selected from the youngest age class (one year old) of mature capsules, and the middle capsule of each umbel (where applicable) was measured. Inflorescence data were collected from up to 70 umbels per tree. Where abortion scars were present on the umbel, or capsules had obviously dislodged during sampling, capsules were scored as present. Leaf measurements were taken on a single representative leaf. Measurements were combined with those from Jones *et al.* (2002).

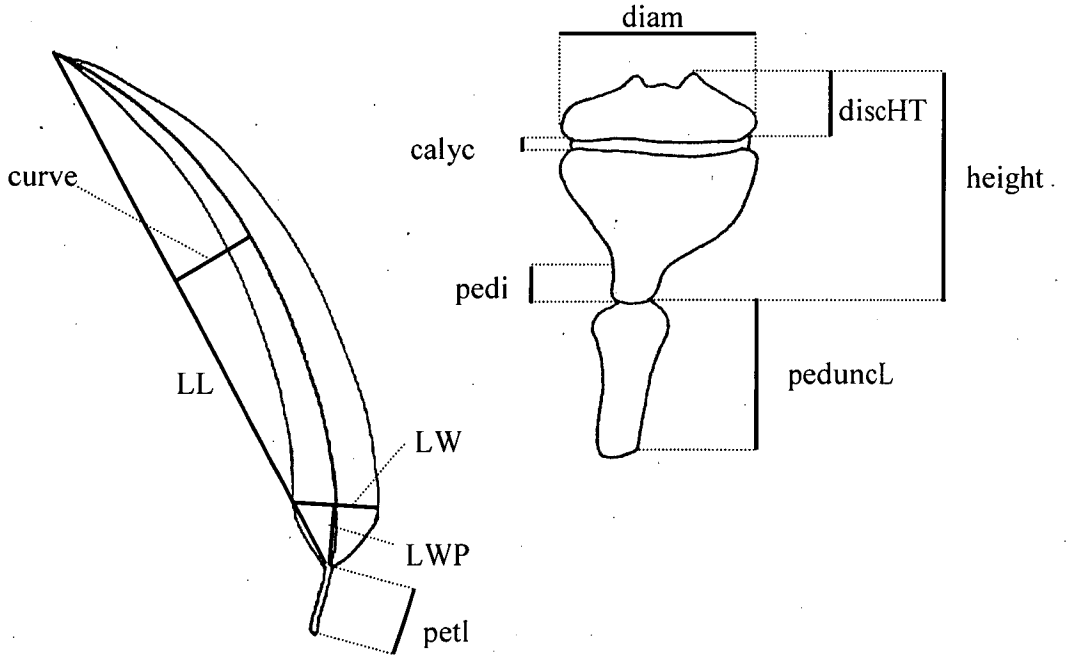


Figure 1.4. Morphological variables measured for the leaves and capsules of individuals of the *E. globulus* complex. Variable codes are described in Table 1.3.

Table 1.3. Morphological variables measured for *E. globulus* in this study. The transformations used to optimise normality and homogenise variances are indicated.

Variable code	Variable description	Transformation
Capsules		
diam	Capsule diameter (mm)	square root
height	Total capsule height (mm)	none
discHT	Disc height (mm)	square root
calyc	Calycine ring height (mm)	square root
ribs	Number of ribs per capsule	none
pedi	Pedicel length (mm)	none
peduncL	Peduncle length (mm)	ln
Inflorescence		
bd_umb	Mean number of buds per umbel	squared
cap_umb	Mean number of capsules per umbel	squared
Leaves		
petl	Petiole length (mm)	ln
lw	Lamina width (mm)	ln
lwp	Length to widest point of leaf (mm)	ln
LL	Total lamina length (mm)	ln
curve	Curvature of leaf (mm)	ln

1.2.3.2 Morphometric data analysis

Analyses were undertaken on the means of the five capsule replicates from each tree. The variables were transformed as in Jones *et al.* (2002) (Table 1.3) to optimise population normality and remove the correlation between locality or region means and variances. The bd_umb and petl variables were excluded from the analysis due to insufficient data.

Multivariate analyses were undertaken at the region level using Canonical Discriminant Analysis, an ordination technique that maximises differences between groups relative to within-group variation, using the PROC DISCRIM procedure in the statistical package

SAS (SAS Institute; Version 9.1). The locality level mean canonical variate scores were placed within the multivariate discriminant space defined by the region level analysis. Analyses were undertaken firstly using all 12 variables, then using the subset of eight capsule/inflorescence variables and finally using the subset of four leaf variables. The vectors indicating the direction of variation in each variable were constructed from the standardised canonical coefficients. The lengths of the vectors were proportional to the ANOVA *F*-values (see Jordan *et al.* 1993), calculated between groups for each variable.

1.2.4 Relationship between morphological, geographic and molecular affinities

1.2.4.1 Relationship between morphological, geographic and molecular affinities of regions across the *E. globulus* species complex

Pairwise Mahalanobis distances among regions and significance ($P<0.05$) were calculated using 12 variables and PROC DISCRIM. A Mantel test was conducted in XLSTAT (Version 2006.5; Addinsoft SARL, Paris, France) to test the relationships between Mahalanobis distance, Nei's (1978) genetic distance (calculated from microsatellite data) and geographic distance (calculated from latitude/longitudes in GenAlEx 6.2, Peakall and Smouse 2006) among regions.

1.2.4.2 Correlations between morphological and molecular affinities in intergrade and geographically outlying populations of the *E. globulus* species complex

Using morphological data, the core *bicostata*, *globulus*, *pseudoglobulus* and *maidenii* individuals (as defined in 1.2.2.2.2, see also Appendix 3) were used to define the multivariate space and the intergrade/geographically outlying individuals were fitted within this space. The morphological affinities (M) of these individuals to each of the core taxa (B, G, P, M) were then calculated using PROC DISCRIM (i.e. M_B, M_G, M_P, M_M). PROC CORR was then used to test whether the morphological affinities of individuals were significantly correlated with their molecular affinities, in each region in Victoria and NSW, using Pearson's and Spearman's correlation coefficients. The molecular affinities had been calculated in the STRUCTURE analysis that used *a priori* population information for the same core individuals, and calculated the affinities of individuals (Q) to each of the four core groups (see 1.2.2.2.2) (i.e. Q_B, Q_G, Q_P, Q_M).

The relationship between morphological (M) and molecular (Q) affinities of individuals to each taxon, for a given region, was graphed (i.e. M_B *vs.* Q_B, M_G *vs.* Q_G, M_P *vs.* Q_P, and M_M *vs.* Q_M). Only regions where there was a significant correlation were graphed, as well as some examples of the individual affinities in a region used to define one of the cores.

Morphological and molecular affinities of individuals were mapped for the following regions: 2-Mt. Cole, 11-Alfred-Nadgee, 14-Lerderderg Gorge, 15-Buchan, 16-Omeo, 17-Mitchell River and 18-Strzelecki Ranges. Each individual was allocated to a morphological and molecular cluster based on the Q-values. Individuals with a Q-value above 0.7 were allocated to their respective cluster, and the remaining individuals were considered admixed. The nature of this admixture was determined by allocating admixed individuals to two or more groups based on a Q-value greater than 0.25.

1.3 Results

1.3.1 Molecular results

1.3.1.1 Microsatellite repeatability

Within this study, there was a 6% error rate (12 mismatches in 200 repeated alleles, data not shown). There was a 3.3% error rate (4 mismatches in 120 alleles) when re-genotyping the set of ten DNA samples from Steane *et al.* (2006) for the six EMCRC loci genotypes (data not shown). Steane (unpubl.) genotyped the set of ten DNA samples for the four EMBRA loci and compared genotypes with those obtained in this study, and reported a 2.5% error rate (2 mismatches in 80 alleles). Thus the overall error rate was 4.5%.

1.3.1.2 Microsatellite variation and differentiation

All nine loci used were highly polymorphic, with 21 - 68 observed alleles per locus (mean A across loci, A_s, = 32, Table 1.4) and high levels of expected heterozygosity (mean H_e across loci, H_T, = 0.89, Table 1.4). The observed heterozygosity was higher than the expected heterozygosity at all loci, and this was reflected in their positive F values, with a mean value of 0.23 (Table 1.4). There was evidence of a null allele at a moderate frequency in most regions in EMBRA19 and EMCRC10 (Table 1.4). After accounting for null alleles, F_{IS} was still greater than zero in the following regions: 9-Wadbilliga, 11-Alfred-Nadgee, 30-SE Tasmania and 32 NE Tasmania (Figure 1.5). As these regions covered a wide geographic range, F_{IS} was calculated for localities within these regions with a sample size greater than 15 (18-Wadbilliga South, 19-Murrabrine, 22-Mumbulla SF, 30-Nadgee, 83-Maria Island South, 91-Cape Tourville Dwarf, 92-Cape Tourville Tall), after accounting for null alleles. In all of these localities, the F_{IS} was lower than in the regional level analysis (Figure 1.5). Pairwise combinations of loci were tested for linkage disequilibrium and none of the loci combinations were significant ($P<0.05$) in all 33 regions.

Table 1.4. Overall genetic diversity parameters for the microsatellite loci used in the analysis of the *E. globulus* species complex.

Locus	*n*	Size range (bp)	*A*	H_e	H_o	*F*	Null
EMBRA11	1148	88-164	35	0.92	0.78	0.15	0.11
EMBRA19	1146	141-201	21	0.81	0.49	0.40	0.12
EMBRA30	1155	81-161	34	0.92	0.79	0.14	0.02
EMCRC2	1132	155-213	26	0.89	0.74	0.17	0.03
EMCRC5	1155	143-351	68	0.88	0.63	0.29	0.08
EMCRC6	1155	148-204	26	0.90	0.70	0.22	0.04
EMCRC7	1132	248-328	29	0.90	0.75	0.16	0.03
EMCRC10	1134	301-355	22	0.87	0.48	0.44	0.17
EMCRC11	1165	217-273	27	0.90	0.81	0.11	0.01
Mean	1147		32	0.89	0.68	0.23	0.06

n = number of individual trees. The number of individual trees varies among loci because clear amplification products were not obtained for all individuals. *A* = observed number of alleles per locus, H_e = expected heterozygosity, H_o = observed heterozygosity, *F* = Wright's Fixation Index. Null = frequency of null alleles, averaged over regions.

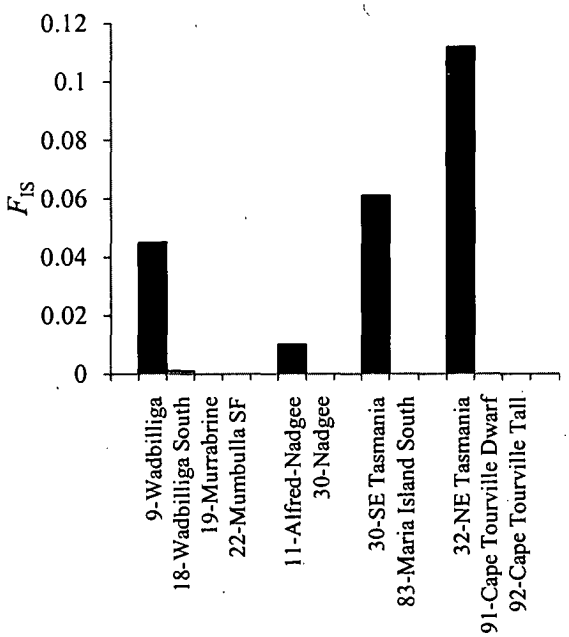


Figure 1.5. F_{IS} for regions (black) of the *E. globulus* complex, followed by localities within these regions (grey/empty). Only regions with a positive F_{IS} after accounting for null alleles are shown, and only localities within these regions with greater than 15 individuals per locality were analysed.

Some geographically disjunct regions of the species complex had lower genetic diversity than the mean for the species complex (e.g. 1-Mt. Bryan H_e = 0.57, 7-Wollemi H_e = 0.68, mean H_e = 0.78, Figure 1.6a, Table 1.5), though some populations in the continuous range also exhibited low levels of genetic diversity (e.g. 26 Wilson's Promontory H_e = 0.65, 25 Tidal River H_e = 0.69, Figure 1.6a, Table 1.5). 1-Mt. Bryan, 7-Wollemi and 21-King Island also exhibited a high level of allelic fixation, as measured by the proportion of loci with an allele greater than 0.5 in frequency ($L_{0.5}$; Table 1.5, Figure 1.6b). The Nadgee locality (locality 30) also had a lower expected heterozygosity and higher level of allelic fixation than adjacent localities of *maidenii* (localities 27-29), *pseudoglobulus* (35-37) and their intergrades (localities 31-34) (Appendix 4).

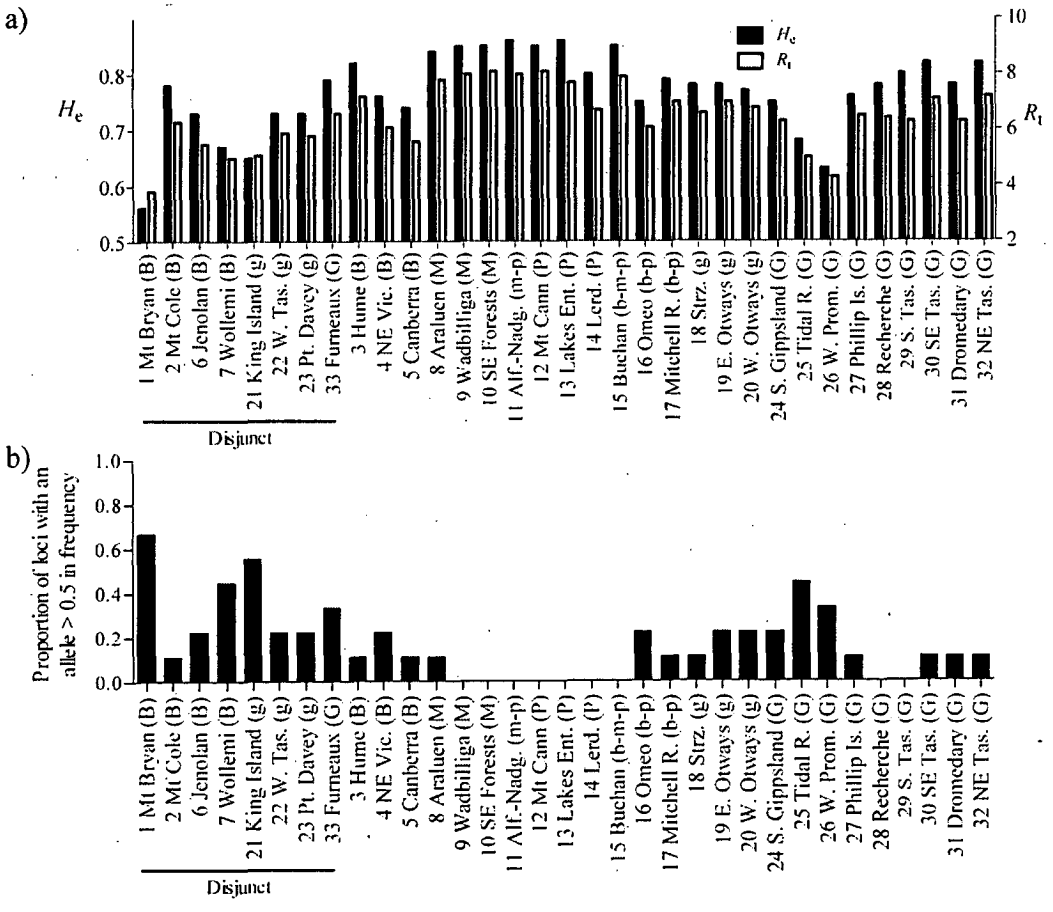


Figure 1.6. a) Expected heterozygosity (H_e) and allelic richness (R_t), and b) proportion of loci with an allele > 0.5 in frequency ($L_{0.5}$), for each region of the *E. globulus* species complex. Disjunct regions (populations isolated from other populations by more than 80 km, and those on the Bass Strait Islands) are indicated; all other regions are in the continuous range of the *E. globulus* distribution.

Table 1.5. Genetic diversity parameters for *E. globulus*, calculated at the regional level.

Region (taxon code)	*n*	*A*	R_t	$L_{0.5}$	H_e	H_o	*F*
1-Mt Bryan (B)	24.0	4.6	3.8	0.7	0.57	0.48	0.17
2-Mt Cole (B)	18.7	8.6	6.3	0.1	0.80	0.67	0.17
3-Hume (B)	29.0	10.8	7.2	0.1	0.84	0.72	0.14
4-NE Victoria (B)	31.2	9.8	6.1	0.2	0.77	0.60	0.23
5-Canberra (B)	19.1	7.3	5.6	0.1	0.76	0.65	0.16
6-Jenolan (B)	12.1	6.2	5.5	0.2	0.76	0.65	0.15
7-Wollemi (B)	19.8	6.8	5.0	0.4	0.68	0.59	0.14
bicostata mean	22.0	7.7	5.6	0.3	0.74	0.62	0.17
8-Araluen (M)	38.6	14.6	7.8	0.1	0.85	0.73	0.14
9-Wadbilliga (M)	112.9	20.2	8.0	0.0	0.85	0.73	0.15
10-SE Forests (M)	59.9	16.4	8.1	0.0	0.86	0.74	0.14
maidenii mean	70.5	17.1	8.0	0.0	0.85	0.73	0.14
11-Alfred-Nadgee (m-p)	53.4	16.0	8.0	0.0	0.87	0.69	0.21
m-p intergrade mean	53.4	16.0	8.0	0.0	0.87	0.69	0.21
12-Mt Cann (P)	25.0	12.7	8.1	0.0	0.86	0.71	0.18
13-Lakes Entrance (P)	19.8	9.3	7.7	0.0	0.82	0.68	0.17
14-Lerderderg Gorge (P)	46.9	13.7	6.7	0.0	0.87	0.73	0.16
pseudoglobulus mean	30.6	11.9	7.5	0.0	0.85	0.71	0.17
15-Buchan (b-m-p)	35.7	13.7	7.9	0.0	0.86	0.69	0.21
b-m-p intergrade mean	35.7	13.7	7.9	0.0	0.86	0.69	0.21
16-Omeo (b-p)	21.9	8.9	6.1	0.2	0.77	0.66	0.15
17-Mitchell River (b-p)	18.8	9.7	7.0	0.1	0.81	0.71	0.12
b-p intergrade mean	20.4	9.3	6.6	0.2	0.79	0.69	0.14
18-Strzelecki Ranges (g)	23.0	9.4	6.6	0.1	0.80	0.70	0.13
19-Eastern Otways (g)	27.3	10.6	7.0	0.2	0.79	0.70	0.12
20-Western Otways (g)	30.4	11.2	6.8	0.2	0.79	0.70	0.11
21-King Island (g)	36.9	7.9	5.1	0.6	0.65	0.56	0.14
22-Western Tasmania (g)	31.0	9.1	5.9	0.2	0.75	0.63	0.16
23-Port Davey (g)	33.8	9.8	5.8	0.2	0.74	0.64	0.13
g intergrade mean	30.4	9.7	6.2	0.3	0.75	0.66	0.13
24-South Gippsland (G)	27.8	9.9	6.3	0.2	0.77	0.68	0.12
25-Tidal River (G)	29.7	7.3	5.0	0.4	0.69	0.76	-0.10
26-Wilson's Promontory (G)	28.6	5.8	4.3	0.3	0.65	0.64	0.00
27-Phillip Island (G)	23.3	10.2	6.5	0.1	0.77	0.71	0.08
28-Recherche Bay (G)	25.4	9.2	6.4	0.0	0.79	0.68	0.15
29-Southern Tasmania (G)	23.9	8.9	6.3	0.0	0.81	0.74	0.10
30-SE Tasmania (G)	51.3	13.6	7.1	0.1	0.82	0.68	0.17
31-Dromedary Tas (G)	34.3	10.8	6.3	0.1	0.79	0.72	0.09
32-NE Tasmania (G)	88.6	14.4	7.2	0.1	0.83	0.69	0.17
33-Furneaux (G)	44.9	11.1	6.6	0.3	0.80	0.69	0.14
globulus mean	37.8	10.1	6.2	0.2	0.77	0.70	0.09
E. globulus mean	34.8	10.6	6.5	0.2	0.78	0.68	0.14

n = mean number of individual trees per locus per population, *A* = mean number of alleles per locus, R_t = allelic richness, $L_{0.5}$= proportion of loci with an allele > 0.5 in frequency, H_e = expected heterozygosity, H_o = observed heterozygosity, *F* = Wright's Fixation Index. For taxon codes see Table 1.2.

Genetic diversity levels in intergrade populations were compared with their putative parental populations, as a recent hybrid zone would be expected to have increased diversity due to gene flow from differentiated populations. However, intergrade regions or localities did not necessarily have higher levels of genetic variation than their putative parental populations. For example, 16-Omeo and 17-Mitchell River, which were classified in the field as intermediate in morphology between *pseudoglobulus* and *bicostata*, had intermediate levels of diversity (mean H_e = 0.79, Table 1.5) compared with pure populations of *pseudoglobulus* (mean H_e = 0.85, Table 1.5) and *bicostata* (H_e = 0.74, Table 1.5). Similarly, the 15-Buchan b-m-p intergrade did not have higher diversity than the closest pure populations of *bicostata* (4-NE Victoria), *maidenii* (10-SE Forests) and *pseudoglobulus* (13-Lakes Entrance) (Table 1.5). However, the mid Cann Valley Highway locality (locality 33) did have marginally lower allelic richness and expected heterozygosity than the localities at the northern (locality 28) and southern (locality 34) ends of the highway (Appendix 4).

A more obvious pattern when comparing genetic diversity between regions was that genetic diversity, as measured by expected heterozygosity and allelic richness, was highest in the SE Forests region, where *pseudoglobulus*, *maidenii* and their intergrades are distributed (regions 8-13, 15; Table 1.5, Figure 1.6a). Allelic fixation was low in this region (Table 1.5, Figure 1.6b). There was a weak but significant negative linear correlation between geographic distance from the SE Forests region (26-Yurammie, 36.9°S, 149.72°E) and genetic diversity measured by both expected heterozygosity and allelic richness, at both the region (Figure 1.7) and locality (Figure 1.8) level. This was driven partly by the large geographic disjunction and low genetic diversity of the Mt. Bryan locality as well as the lower diversity in Tasmania. However when the Mt Bryan population and Tasmanian samples were excluded, the correlation was still significant both at the region (Figure 1.9) and locality (Figure 1.10) level. As neither the distribution of *E. globulus* nor its proposed migration route are linear, plotting these genetic diversity statistics for each region or locality as a contour map gave an improved depiction of areas of high genetic diversity in SE Forests, on the east coast of Tasmania, and on the western side of the Great Dividing Range (Figure 1.11, Figure 1.12).

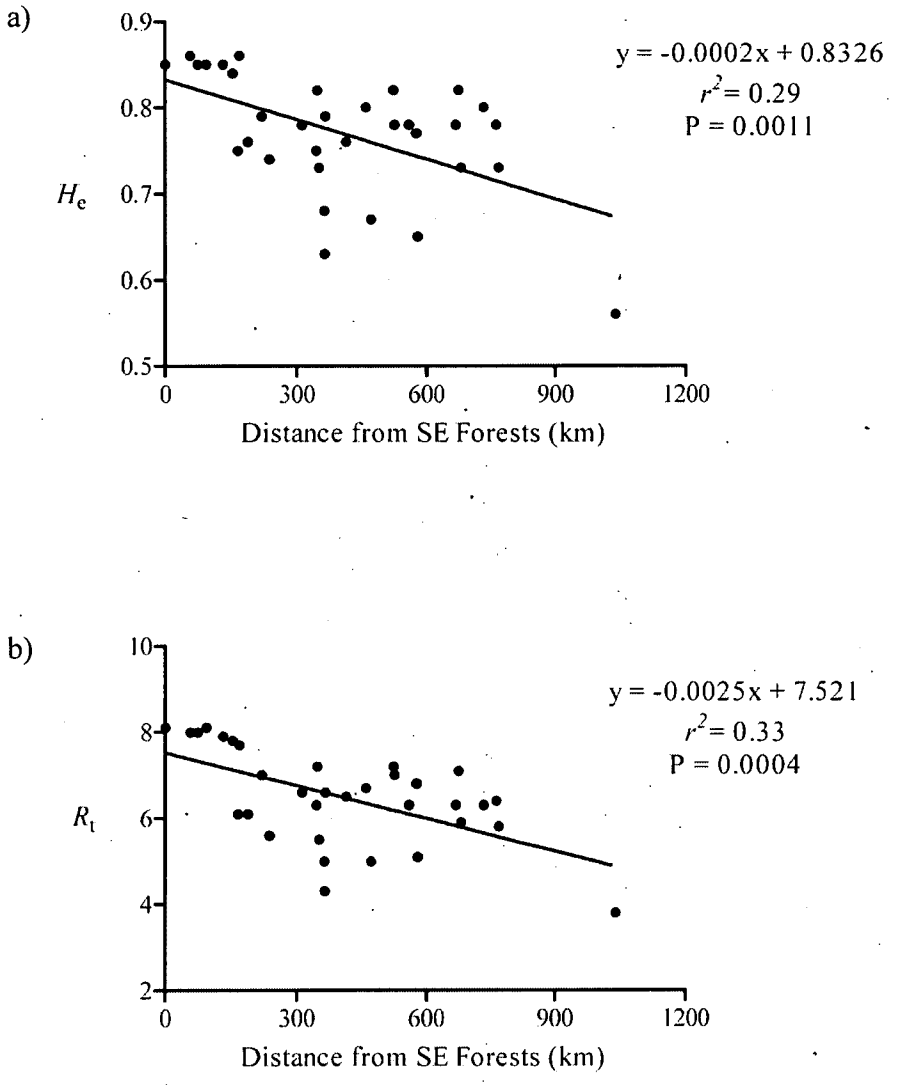


Figure 1.7 Relationship between genetic diversity per region and geographic distance from the SE Forests region for the *E. globulus* species complex. a) expected heterozygosity (H_e) *versus* geographic distance from locality 26 in the SE Forests region. b) allelic richness (R_t) *versus* geographic distance from SE Forests.

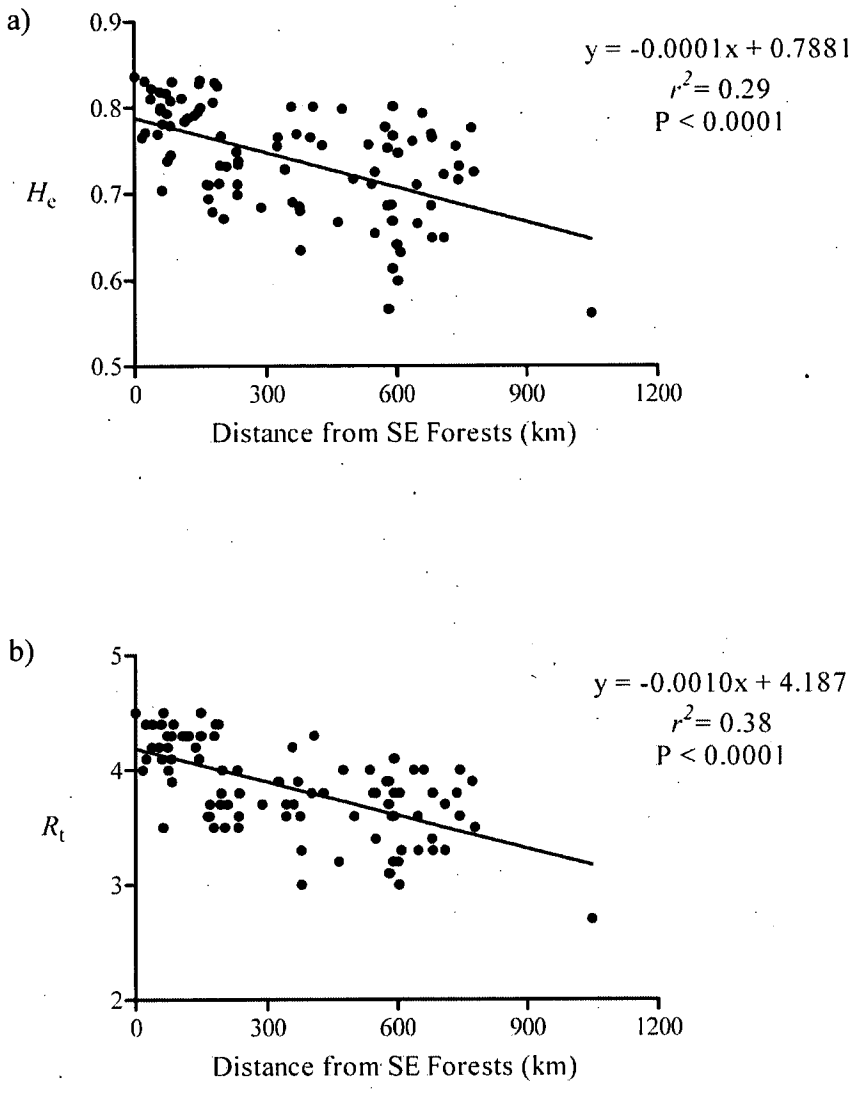


Figure 1.8. Relationship between genetic diversity per locality and geographic distance from the SE Forests region for the *E. globulus* species complex. a) expected heterozygosity (H_e) *versus* geographic distance from locality 26 in the SE Forests region. b) allelic richness (R_t) *versus* geographic distance from SE Forests.

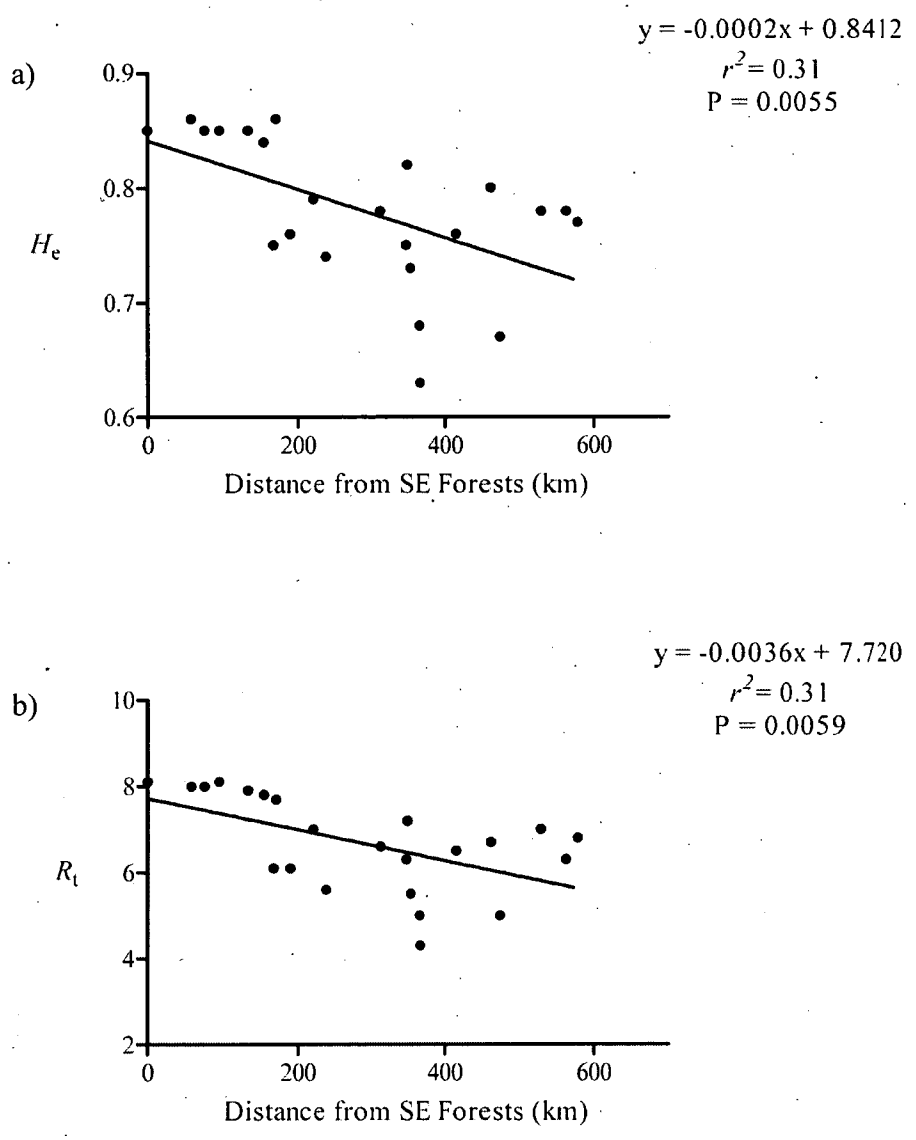


Figure 1.9. Relationship between genetic diversity per region and geographic distance from the SE Forests region for a subset of the *E. globulus* species complex, excluding Tasmanian and South Australian samples. a) expected heterozygosity (H_e) *versus* geographic distance from locality 26 in the SE Forests region. b) allelic richness (R_t) *versus* geographic distance from SE Forests.

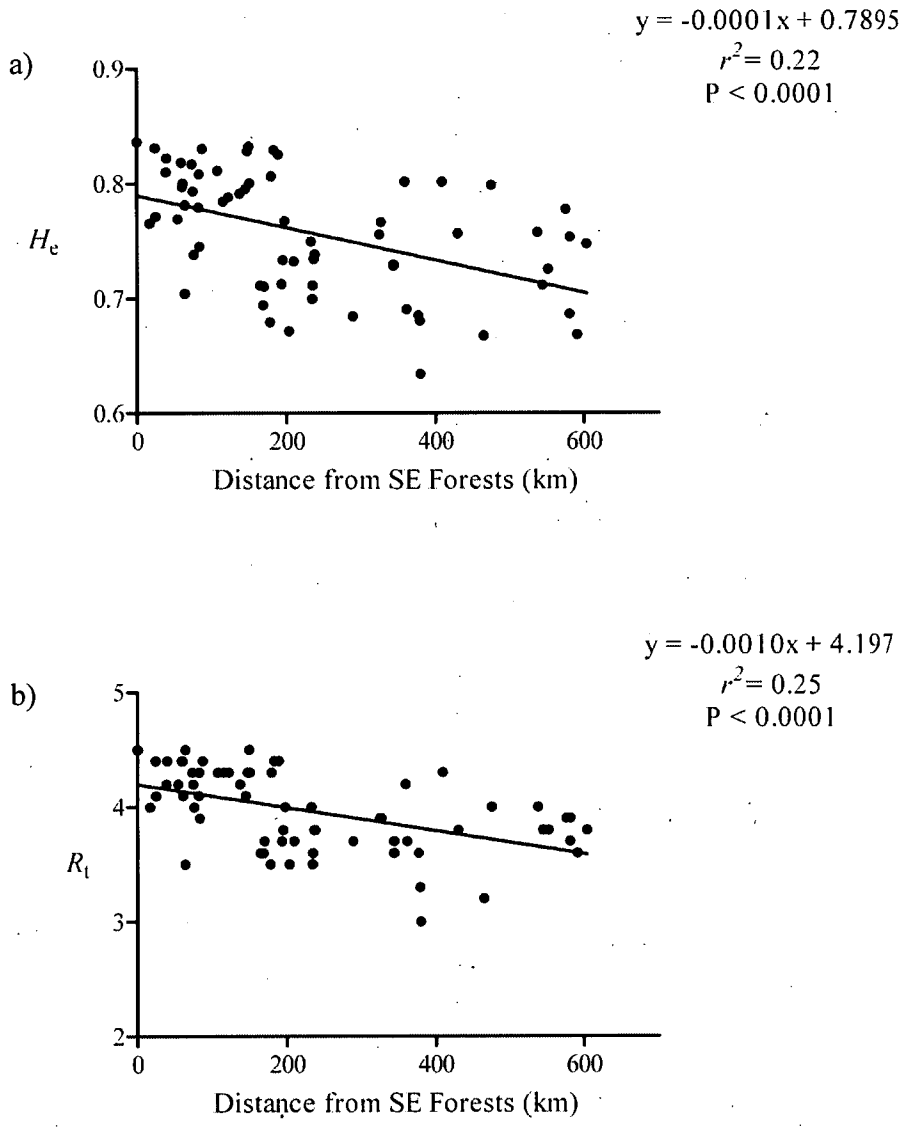


Figure 1.10. Relationship between genetic diversity per locality and geographic distance from the SE Forests region for a subset of the *E. globulus* species complex, excluding Tasmanian and South Australian samples. a) expected heterozygosity (H_e) *versus* geographic distance from locality 26 in the SE Forests region. b) allelic richness (R_t) *versus* geographic distance from SE Forests.

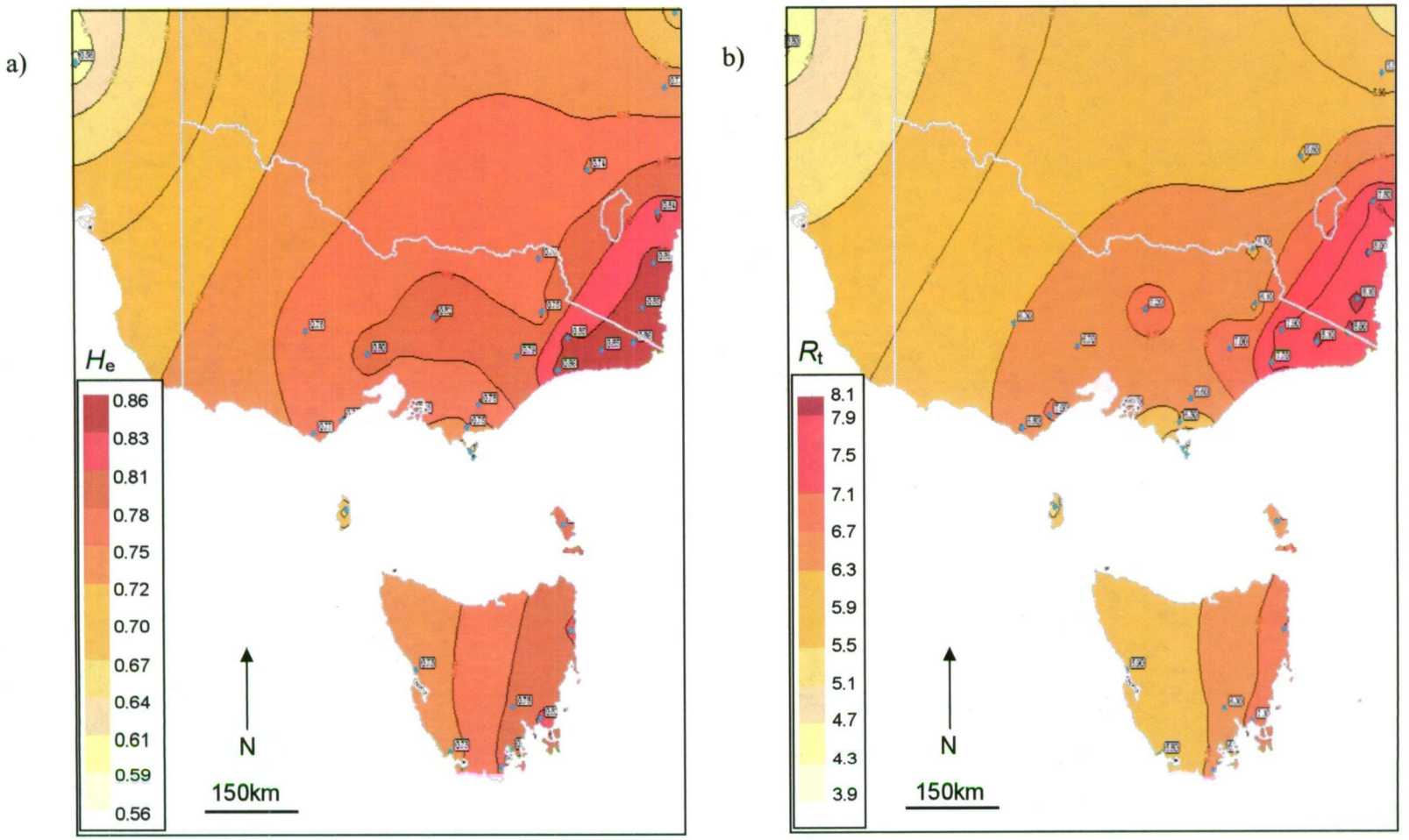


Figure 1.11. Genetic diversity statistics for regions of *E. globulus*, plotted as a contour map. a) expected heterozygosity (H_e) and b) allelic richness (R_t) per region.

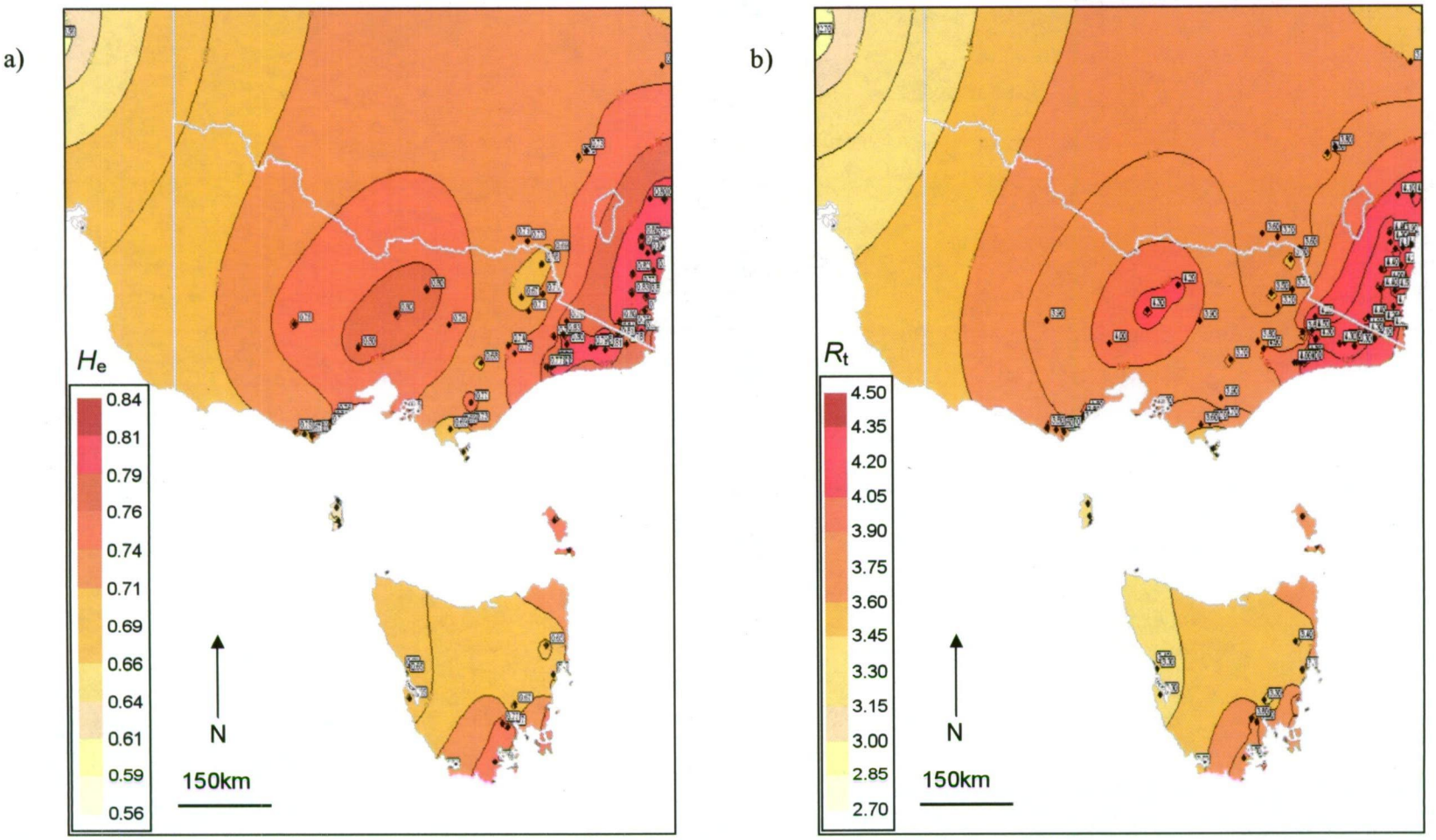


Figure 1.12. Genetic diversity statistics for localities of *E. globulus*, plotted as a contour map. a) expected heterozygosity (H_e) and b) allelic richness (R_t) per locality.

1.3.1.3 Population differentiation in the *E. globulus* species complex

1.3.1.3.1 Overall differentiation between localities and regions

Most of the genetic variation was between individuals within localities or regions, however there was a significant proportion of genetic variation maintained between localities or regions. Considering locality as the subpopulation and ignoring regional structure, overall F_{ST} (i.e. inbreeding within a locality relative to the entire species complex) = 0.12 and D_{est} = 0.49 (Table 1.6). Using the region as the subpopulation, overall F_{ST} (i.e. inbreeding within a region relative to the entire species complex) = 0.10 and D_{est} = 0.48 (Table 1.6). F_{IS} was high both in localities and in regions for EMBRA19 and EMCRC10 (Table 1.6), just as these loci showed a high F (Table 1.4). F_{ST} and D_{est} measures were similar among loci though (Table 1.6), indicating that no single locus was most informative in differentiating localities or regions.

Table 1.6. Inbreeding coefficients (F_{IS}, F_{IT}, F_{ST}) and Jost's (2008) estimated D (D_{est}) calculated for each locus at the locality or regional level, based on the 94 localities or the 33 regions of the *E. globulus* complex (see Table 1.2). The 95% confidence intervals on the mean inbreeding coefficients, derived from 1000 bootstraps, are given in parentheses.

	Locality level				Regional level			
Locus	F_{IS}	F_{IT}	F_{ST}	D_{est}	F_{IS}	F_{IT}	F_{ST}	D_{est}
EMBRA11	0.06	0.15	0.10	0.50	0.08	0.15	0.08	0.49
EMBRA19	0.31	0.40	0.14	0.40	0.32	0.40	0.12	0.38
EMBRA30	0.04	0.14	0.11	0.53	0.05	0.15	0.10	0.56
EMCRC2	0.06	0.17	0.12	0.45	0.08	0.17	0.10	0.45
EMCRC5	0.19	0.29	0.13	0.51	0.21	0.29	0.11	0.53
EMCRC6	0.08	0.22	0.15	0.58	0.11	0.22	0.13	0.58
EMCRC7	0.05	0.17	0.12	0.50	0.07	0.17	0.10	0.47
EMCRC10	0.36	0.44	0.13	0.48	0.38	0.45	0.10	0.43
EMCRC11	0.02	0.11	0.09	0.46	0.03	0.11	0.08	0.49
Mean	0.12	0.23	0.12	0.49	0.14	0.23	0.10	0.48
	(0.06-0.21)	(0.16-0.31)	(0.11-0.13)		(0.07-0.23)	(0.16-0.31)	(0.09-0.11)	

1.3.1.3.2 Differentiation within and among core taxa of the E. globulus *species complex*

Within the subset of "core" samples used for the AMOVA, differences among taxa accounted for 7% of the genetic variance ($P < 0.001$, Figure 1.13). The difference among regions within taxa accounted for 3% of the variance ($P < 0.001$), while 4% of the variance was among localities within regions ($P < 0.001$, Figure 1.13). The largest proportion of the variance (86%) was among individuals within localities (Figure 1.13).

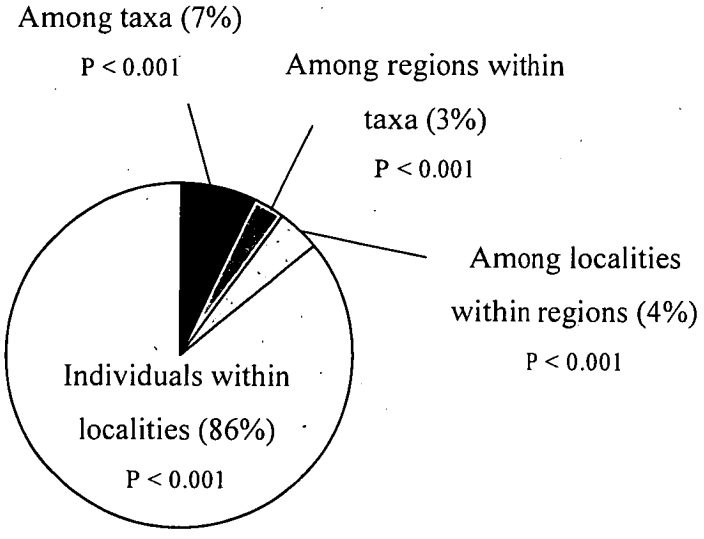


Figure 1.13. Analysis of molecular variance (AMOVA) of morphological "core" samples of the *E. globulus* species complex (as listed in Appendix 1), showing the hierarchical partitioning of genetic variation among populations, regions and taxa.

F_{ST} and D_{est} measures calculated within each morphological taxon were variable among taxa (Table 1.7). Again, most of the genetic variation was between individuals within localities or regions, and there was a significant proportion of genetic variation maintained between localities or regions. Each taxon was ranked similarly in terms of differentiation, whether the F_{ST} or D_{est} statistic was used. Differentiation was lowest within *maidenii* (F_{ST} = 0.025, D_{est} = 0.177) and highest within *bicostata* (F_{ST} = 0.138, D_{est} = 0.412) considering locality as the subpopulation and ignoring regional structure, and the same rank order of differentiation was observed when using the region as the subpopulation (Table 1.7).

Table 1.7. F_{ST} and D_{est} calculated for each taxon of *E. globulus* at the locality or regional level, using only localities or regions with pure morphology as defined in Table 1.2. In the case of *globulus*, calculations are also given with its intergrades included. The 95% confidence intervals on the mean F_{ST}, derived from 1000 bootstraps, are given in parentheses. n = number of localities or regions.

	Locality level			Regional level		
Taxon	n	F_{ST}	D_{est}	n	F_{ST}	D_{est}
bicostata	13	0.138 (0.120-0.160)	0.412	7	0.140 (0.120-0.165)	0.522
globulus and g intergrades	41	0.100 (0.088-0.114)	0.350	16	0.081 (0.072-0.093)	0.328
globulus	25	0.093 (0.079-0.106)	0.341	10	0.072 (0.059-0.084)	0.327
pseudoglobulus	9	0.039 (0.022-0.061)	0.227	3	0.038 (0.024-0.054)	0.242
maidenii	16	0.025 (0.019-0.032)	0.177	3	0.011 (0.005-0.020)	0.081

Pairwise F_{ST} values revealed that *pseudoglobulus* was the least differentiated taxon (pairwise F_{ST} with other taxa 0.049 – 0.055, Table 1.8). Tasmanian and Victorian core *globulus* had the lowest pairwise F_{ST} and genetic distance (0.038 and 0.201 respectively, Table 1.8).

Table 1.8. Pairwise F_{ST} (below diagonal) and Nei's (1978) genetic distance among pure taxa of the *E. globulus* species complex (defined in Table 1.2). The 95% confidence intervals on the mean F_{ST}, derived from 1000 bootstraps, are given in parentheses. B, *bicostata*; M, *maidenii*; P, *pseudoglobulus*; G, *globulus*.

	B	M	P	G (Tas.)	G (Vic.)
B		0.549	0.407	0.812	0.831
M	0.071 (0.048-0.097)		0.458	0.703	0.609
P	0.055 (0.033-0.080)	0.052 (0.034-0.073)		0.357	0.310
G (Tas.)	0.099 (0.059-0.151)	0.082 (0.056-0.110)	0.049 (0.034-0.068)		0.201
G (Vic.)	0.112 (0.070-0.158)	0.083 (0.048-0.122)	0.050 (0.029-0.083)	0.038 (0.021-0.058)	

1.3.1.4 Genetic affinities of regions of the *E. globulus* species complex

1.3.1.4.1 Genetic relationships among populations of the entire E. globulus *species complex*

Regions of *globulus* and its intergrades clustered tightly together in the Neighbour-Joining radial tree, distinct from the *maidenii* cluster and the *bicostata* cluster (Figure 1.14). Populations from the SE Forests region, including *pseudoglobulus*, *maidenii* and their intergrades, were in the centre of the tree (Regions 8-13 and 15, Figure 1.14). Disjunct populations of *bicostata* (1-Mt. Bryan and 7-Wollemi) were on long branches, as was the population at 26-Wilson's Promontory (Figure 1.14). 6-Jenolan, a region classified as *bicostata*, was on the *maidenii* branch (Figure 1.14).

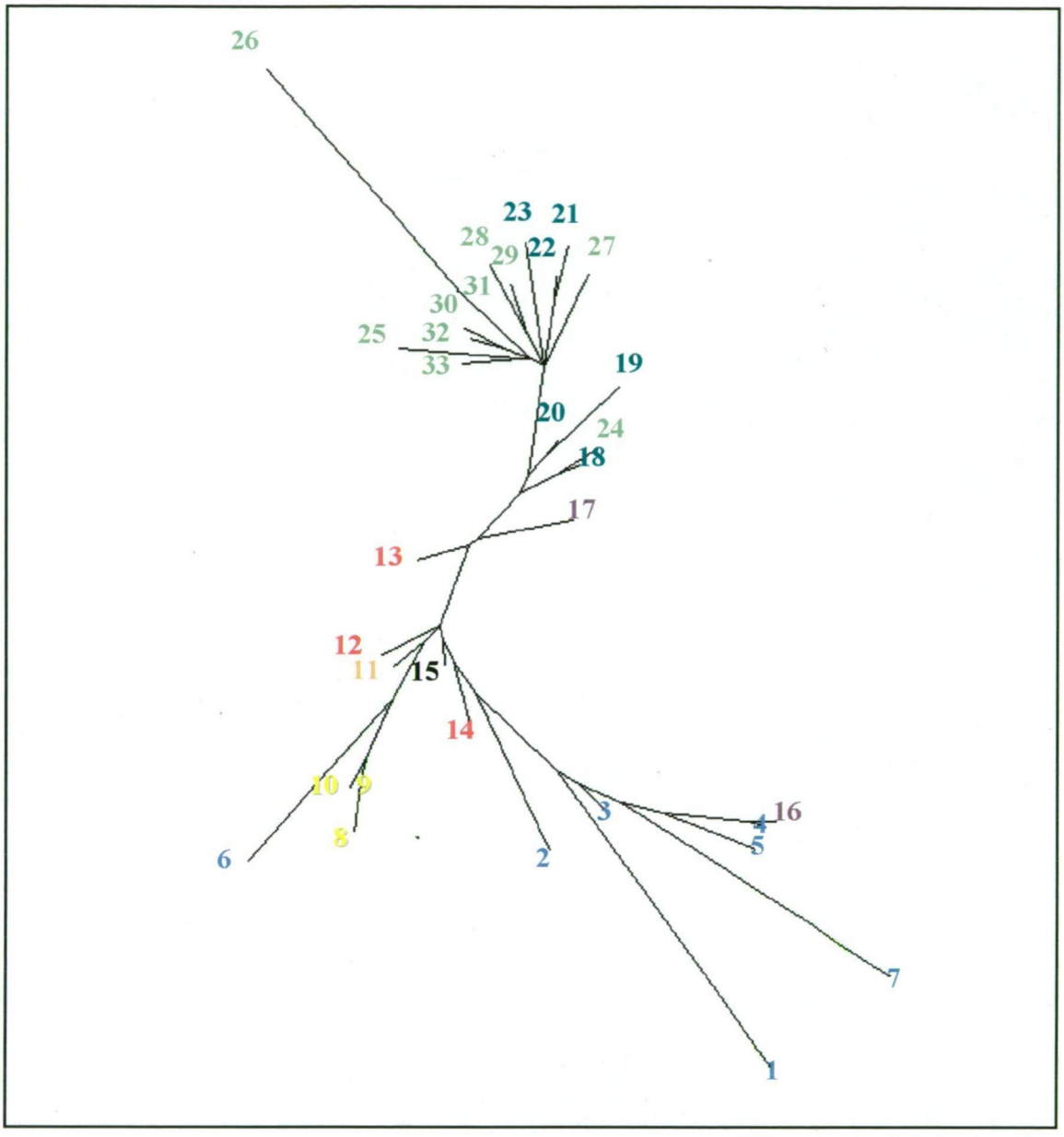


Figure 1.14. Neighbour-Joining radial tree with proportional branch lengths, based on a pairwise matrix of Nei's (1972) genetic distance between regions of the *E. globulus* species complex. For definition of region codes see Table 1.2. Colours of region codes correspond to the observed morphology at the region as in Figure 1.1: B = blue, M = yellow, m-p = orange, P = red, b-m-p = black, b-p = purple, g = dark green, G = pale green.

The ΔK method of Evanno *et al.* (2005) detected the major split in the entire *E. globulus* gene pool at $K = 2$ (Figure 1.15), corresponding to a split between Tasmanian and Victorian individuals, but with populations consisting of admixed individuals in central Victoria linking these two groups (Figure 1.16). At each increase in K the group of regions that split into a new cluster were geographically contiguous (Figure 1.17, Figure 1.18, Figure 1.19, Figure 1.20, Figure 1.21, Figure 1.22), but at each given K there were intermediate individuals that linked the K clusters (Figure 1.16). The intergrade population at Omeo (16) was part of the *bicostata* gene pool at each increase in K (Figure 1.16) though it was classified as intergrade in morphology in the field. Within each region, individuals had similar Q values, i.e. similar levels of admixture and similar affinities to each of the K groups (Figure 1.16). Exceptions, however, were the 11-Alfred-Nadgee and 32-NE Tasmania regions, which consisted of individuals / localities of different molecular affinities (Figure 1.16).

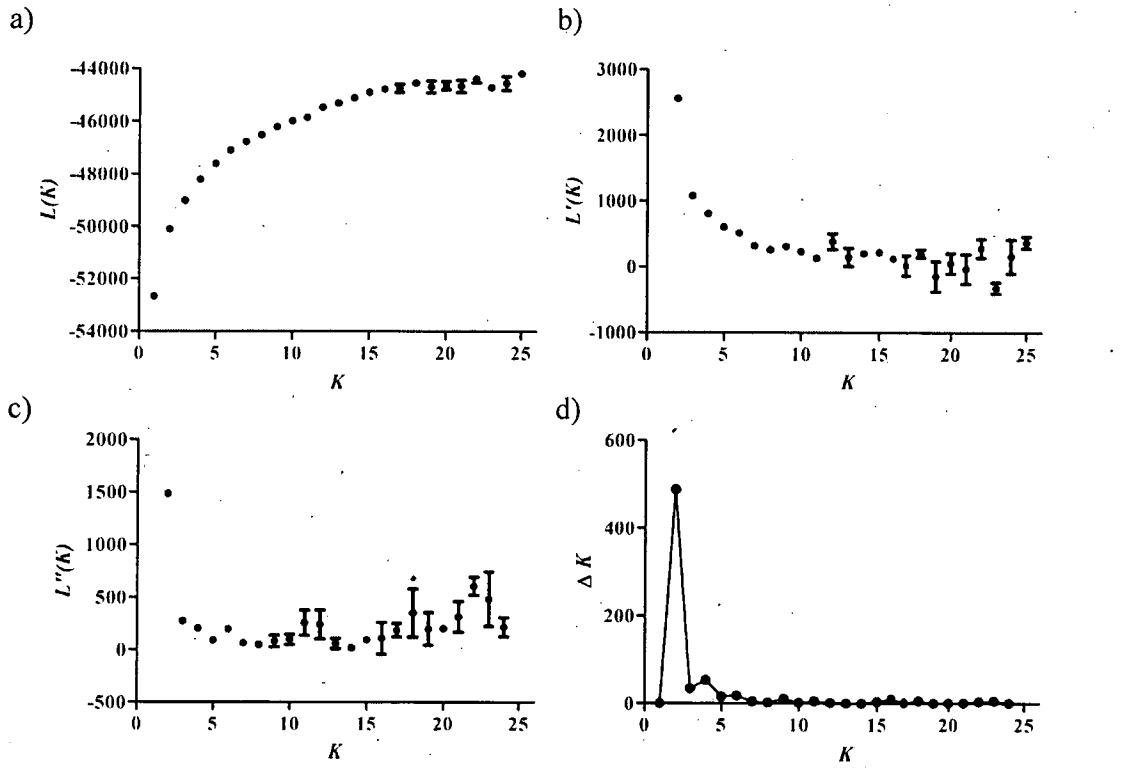


Figure 1.15. The four steps used for the calculation of ΔK for the full *E. globulus* data set (1194 individuals, 33 regions), following the calculations of Evanno *et al.* (2005). a) Mean *L(K)* (± SD) over three highest likelihood runs. b) Mean rate of change of the likelihood distribution *L'(K)* (± SD). c) Mean absolute values of the second order rate of change of the likelihood distribution *L''(K)* (± SD). d) ΔK calculated as the mean in c) divided by the SD in a). The modal value indicates the "true" *K*, or number of clusters; in this case, *K*=2.

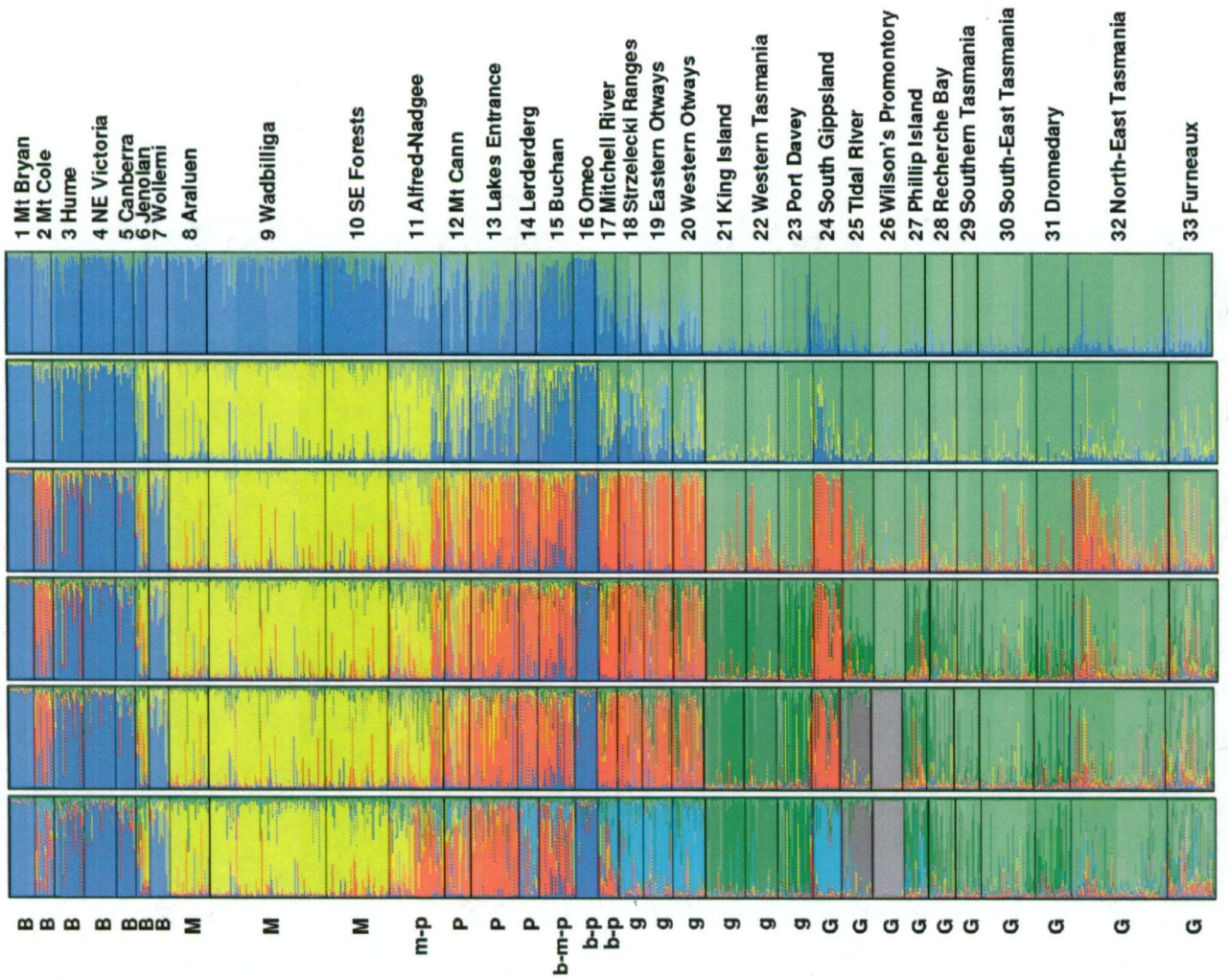


Figure 1.16. Proportion of membership for all sampled individuals of the *E. globulus* species complex. Each individual is represented by a thin vertical line, partitioned into *K* coloured segments representing the individual's estimated membership into the *K* genetic clusters. Individuals are grouped by region/race (indicated above the figure), and the capsule morphology of each region is shown below the figure.

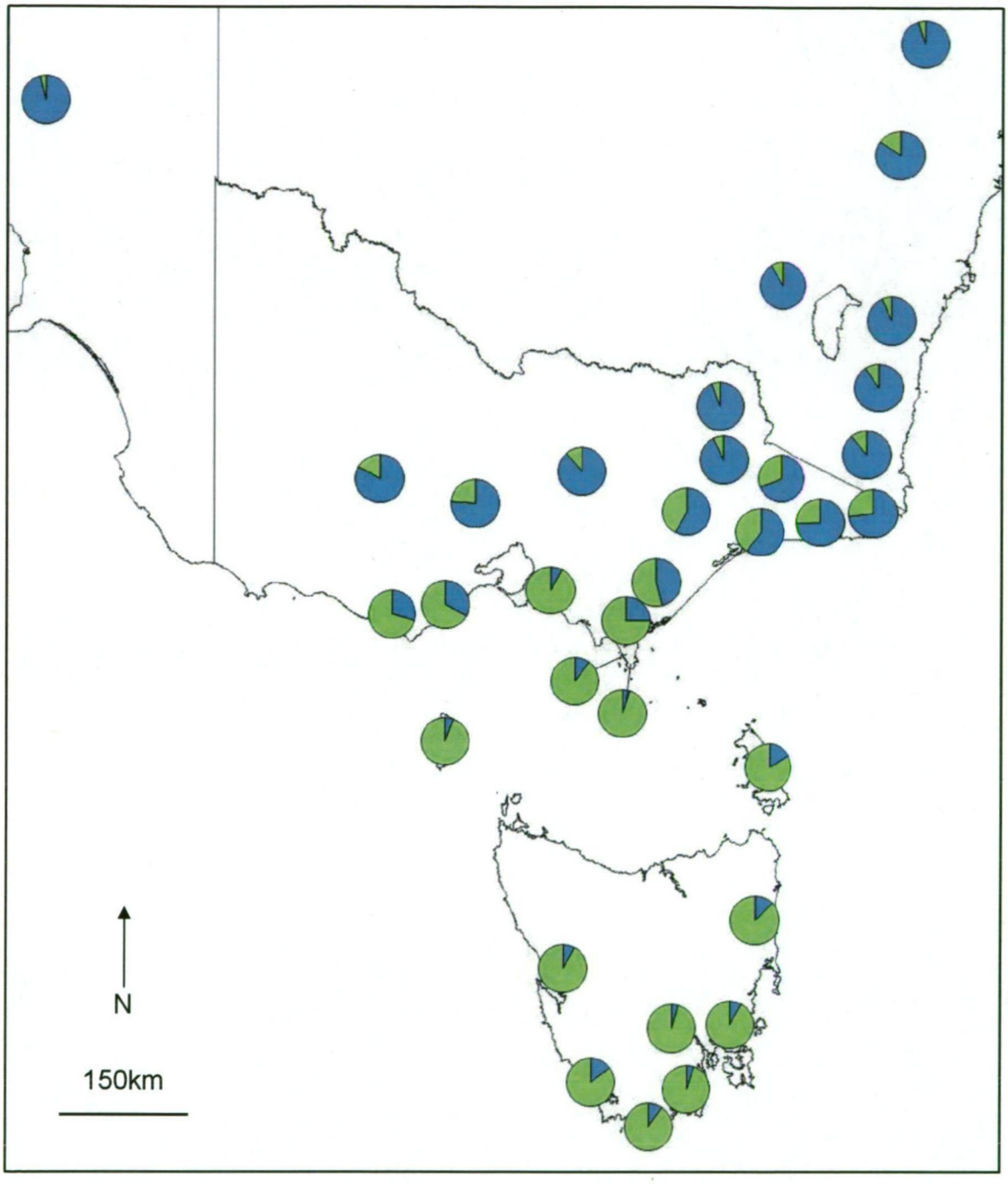


Figure 1.17. The average proportion of membership of each of the 33 regions into each of the clusters ($K = 2$), from the analysis of all sampled individuals of the *E. globulus* complex. Morphological affinities of regions are detailed in Table 1.2 and Figure 1.2.

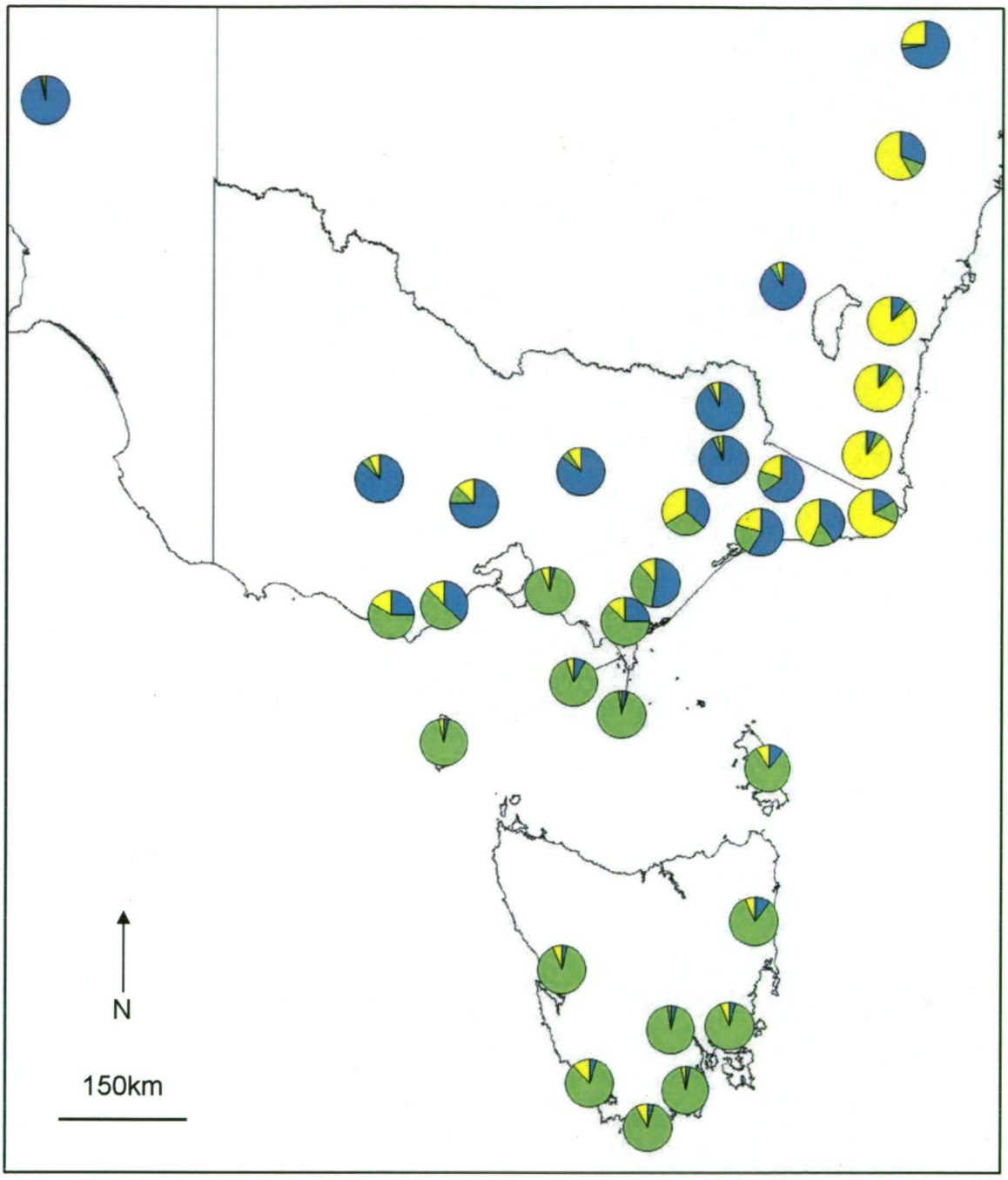


Figure 1.18. The average proportion of membership of each of the 33 regions into each of the clusters ($K = 3$), from the analysis of all sampled individuals of the *E. globulus* complex. Morphological affinities of regions are detailed in Table 1.2 and Figure 1.2.

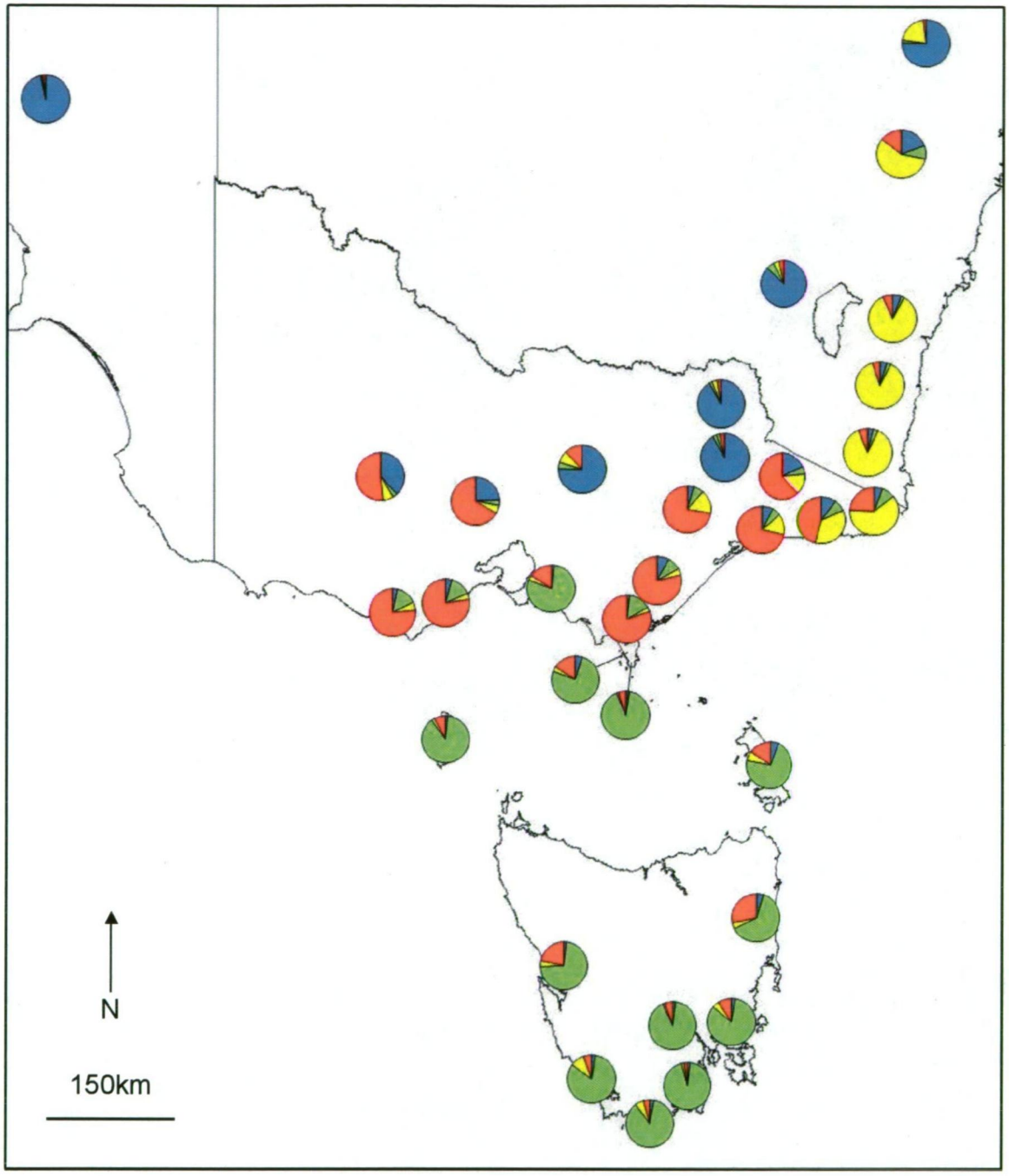


Figure 1.19. The average proportion of membership of each of the 33 regions into each of the clusters (K = 4), from the analysis of all sampled individuals of the *E. globulus* complex. Morphological affinities of regions are detailed in Table 1.2 and Figure 1.2.

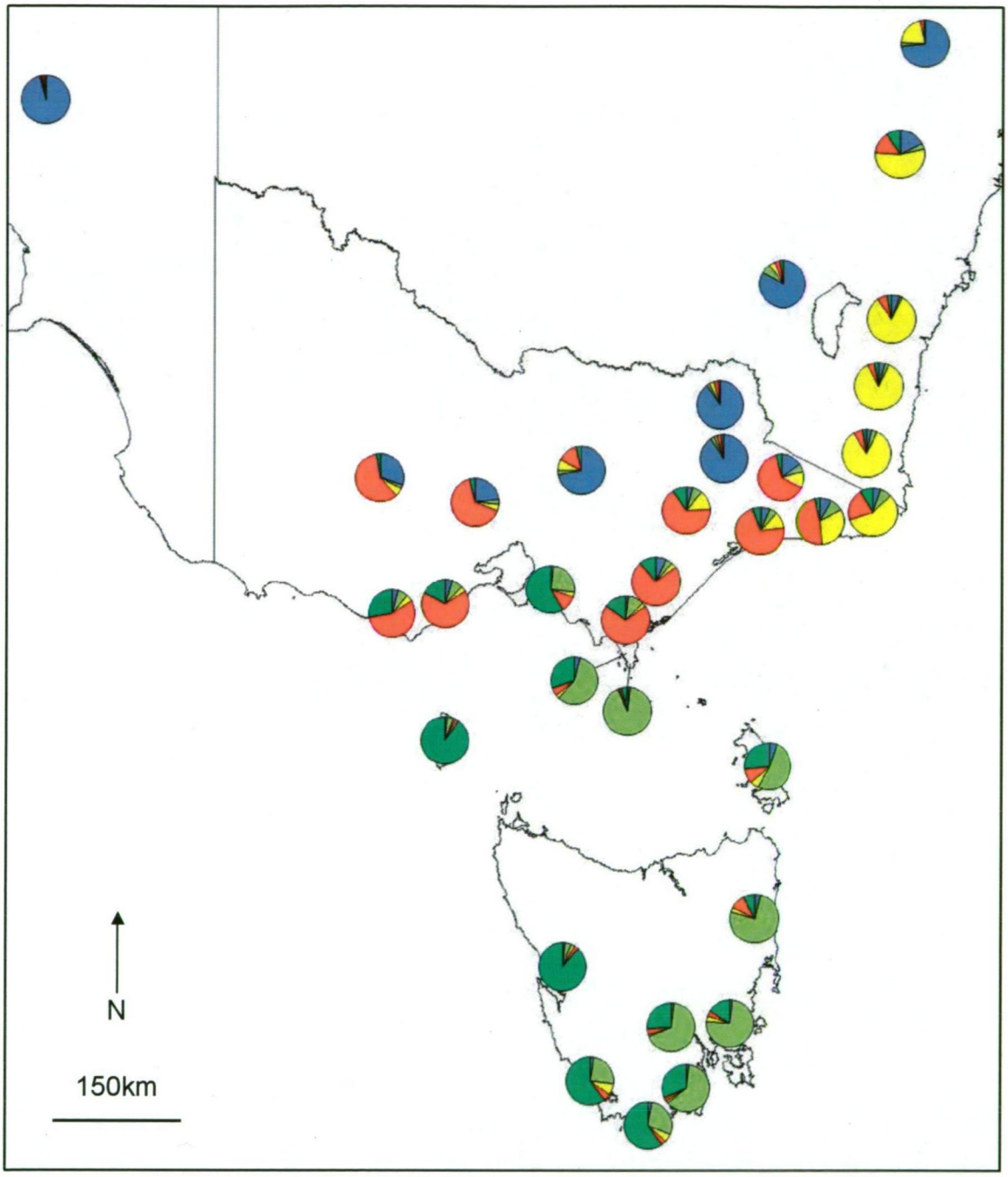


Figure 1.20. The average proportion of membership of each of the 33 regions into each of the clusters (K = 5), from the analysis of all sampled individuals of the *E. globulus* complex. Morphological affinities of regions are detailed in Table 1.2 and Figure 1.2.

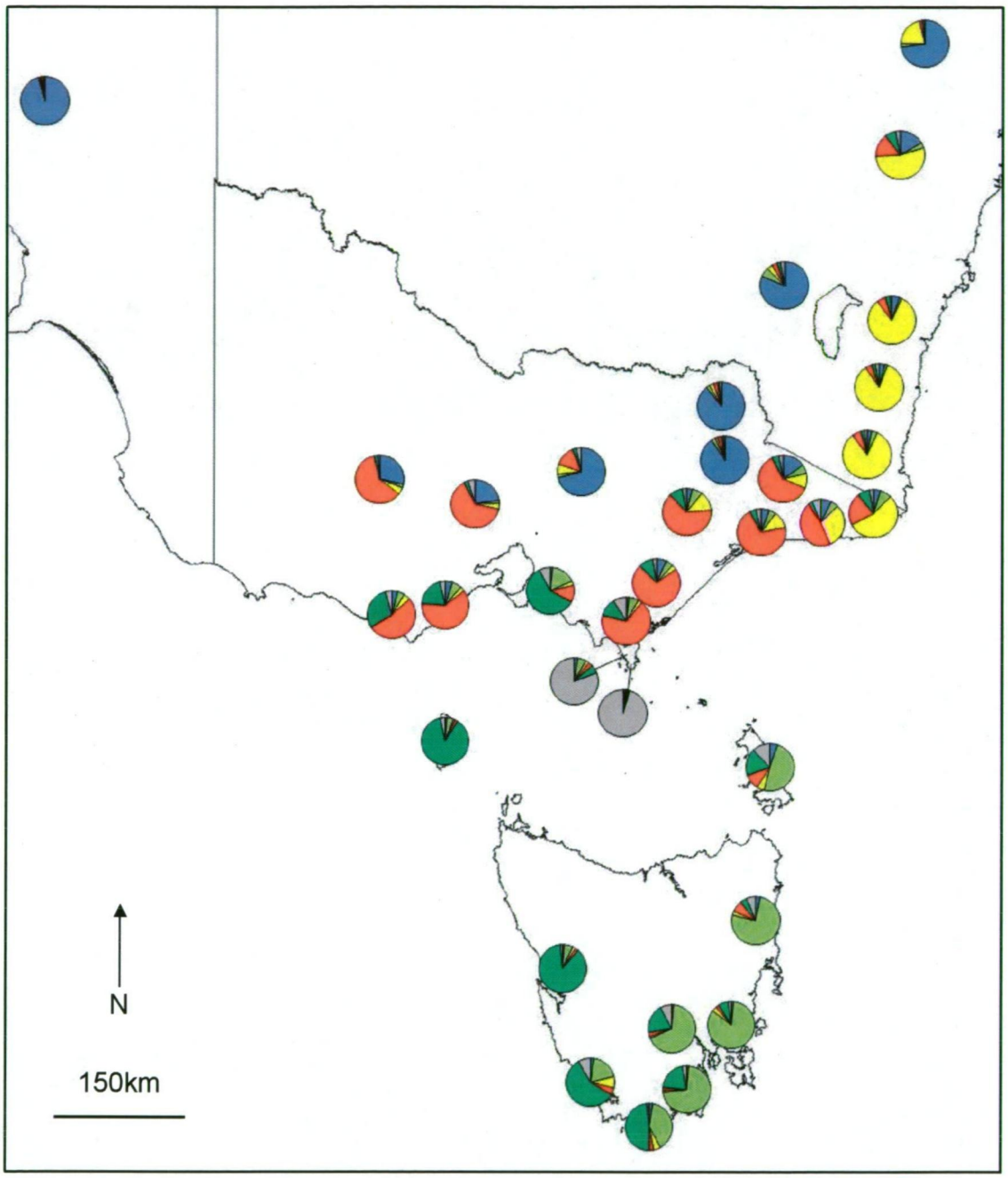


Figure 1.21. The average proportion of membership of each of the 33 regions into each of the clusters ($K = 6$), from the analysis of all sampled individuals of the *E. globulus* complex. Morphological affinities of regions are detailed in Table 1.2 and Figure 1.2.

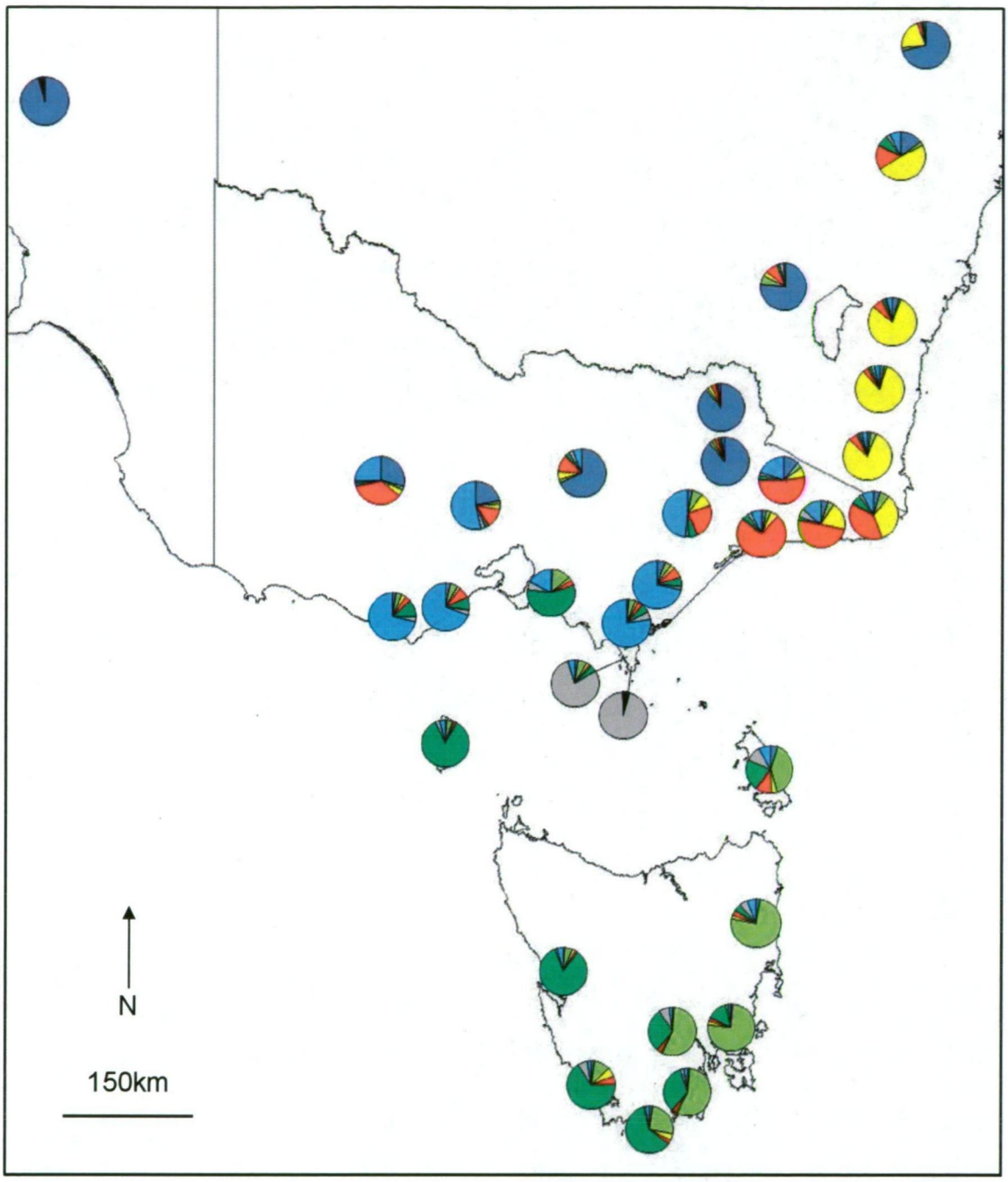


Figure 1.22. The average proportion of membership of each of the 33 regions into each of the clusters (K = 7), from the analysis of all sampled individuals of the *E. globulus* complex. Morphological affinities of regions are detailed in Table 1.2 and Figure 1.2.

1.3.1.4.2 Genetic affinities of intergrade populations of the E. globulus *species complex*

In the second set of STRUCTURE analyses, designed to examine the affinities of the intergrade populations of mainland Australia, all Tasmanian populations and geographical outliers on mainland Australia were excluded from the full data set, resulting in a data set of 749 individuals representing 21 regions. This subset of samples split into either two or three gene pools (Figure 1.23d). At $K = 3$, this corresponded to *bicostata*, *maidenii* and *globulus* forming a cluster each, with intermediate individuals in central Victoria, including morphologically core *pseudoglobulus* populations, linking these clusters (Figure 1.24). At $K = 3$, *pseudoglobulus* (regions 12-14) was a mixed taxon (Figure 1.24). Populations of *pseudoglobulus* in East Gippsland consisted of individuals with strong or admixed molecular affinities to *bicostata* in the locality at 42-Metung (west), to *globulus* in 40-Lakes Entrance, 38-Ostlers Road and 39-Lake Tyres (central) and to *maidenii* at 35-Mt. Cann, in the eastern most range of the distribution of this taxon (data not shown). The variation in molecular affinities across the range of this taxon was not in agreement with the morphological affinities of individuals, nearly all of which were pure *pseudoglobulus* (see 1.3.2).

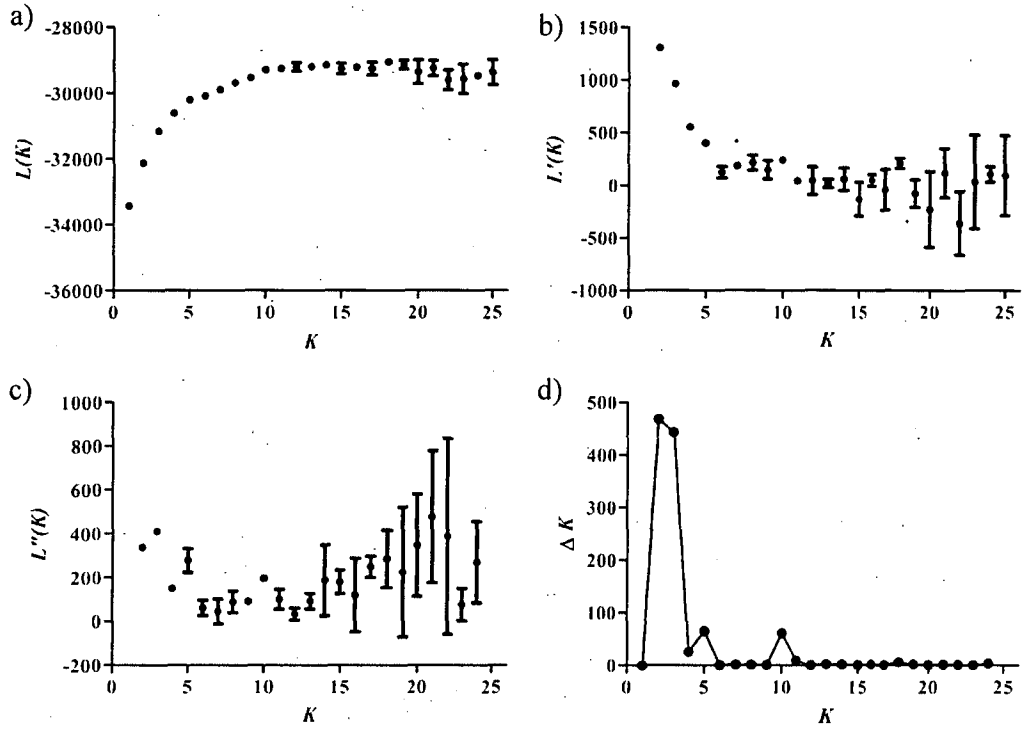


Figure 1.23. The four steps used for the calculation of ΔK for the subset of *E. globulus* samples (749 individuals from 21 regions, see Appendix 2; assuming the admixture model with no *a priori* population groupings) following the calculations of Evanno *et al.* (2005). a) Mean *L(K)* (± SD) over three highest likelihood runs. b) *L'(K)*: mean rate of change of the likelihood distribution (± SD). c) *L''(K)*: mean absolute values of the second order rate of change of the likelihood distribution (± SD). d) ΔK: calculated as the mean in c) divided by the SD in a). The modal value indicates the "true" *K*, or number of clusters; in this case, *K*=2 or 3.

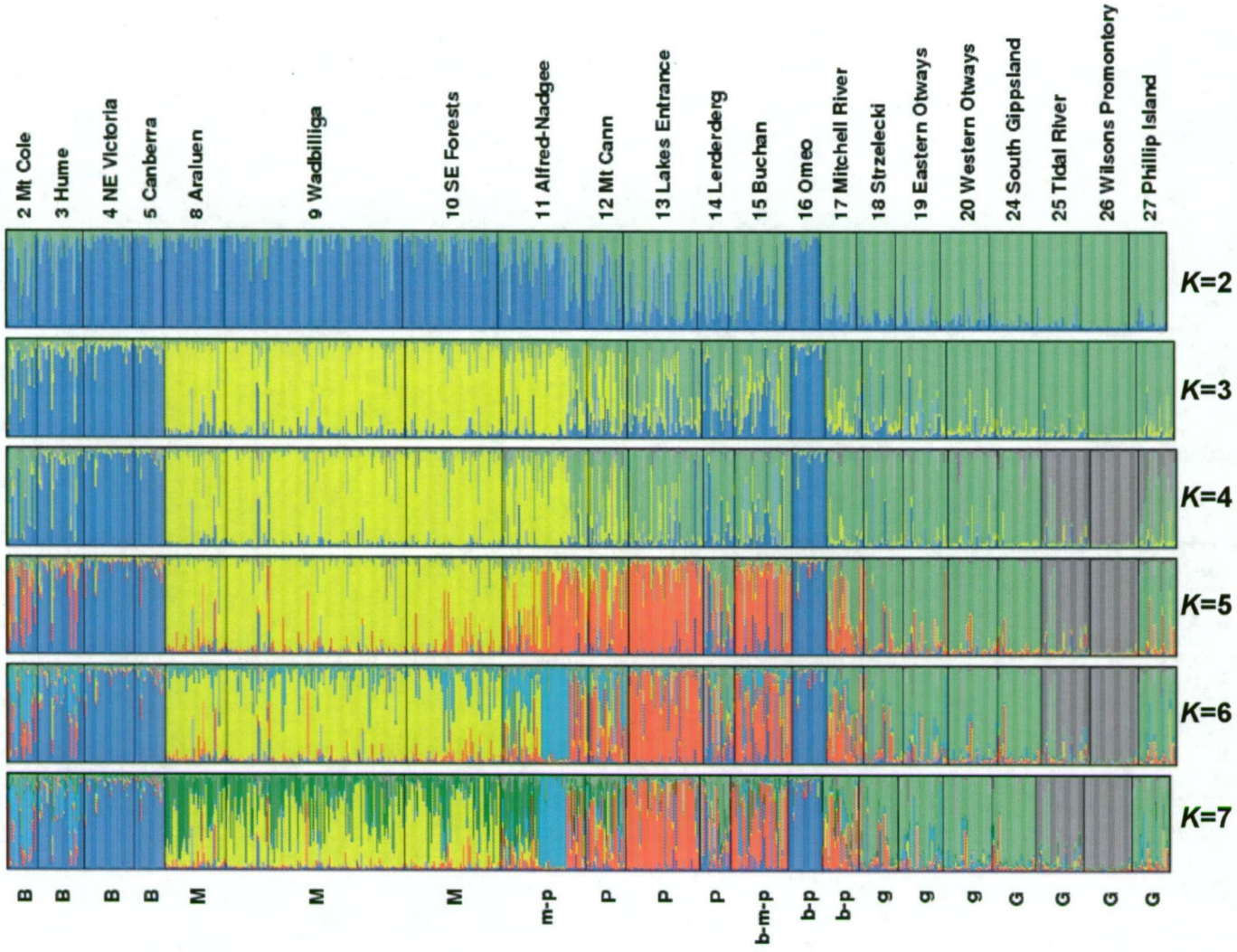


Figure 1.24. Proportion of membership for subset of 749 individuals of the *E. globulus* species complex, (see Appendix 2) using an admixture model with no *a priori* population groupings. Each individual is represented by a thin vertical line, partitioned into *K* coloured segments representing the individual's estimated membership into the *K* genetic clusters. Individuals are grouped by region/race (indicated above the figure), and the capsule morphology of each region is shown below the figure.

1.3.2 Morphological relationships in the E. globulus *species complex*

1.3.2.1 Key morphological characters used to differentiate the regions of the *E. globulus* species complex

Capsule diameter, calycine ring height, number of capsules per umbel and disc height were the most important variables in differentiating the regions of the *E. globulus* complex (Figure 1.25). The analysis using all variables, or only capsule variables, showed the greatest differentiation among taxa, compared with only using leaf variables (Figure 1.25, Figure 1.26).

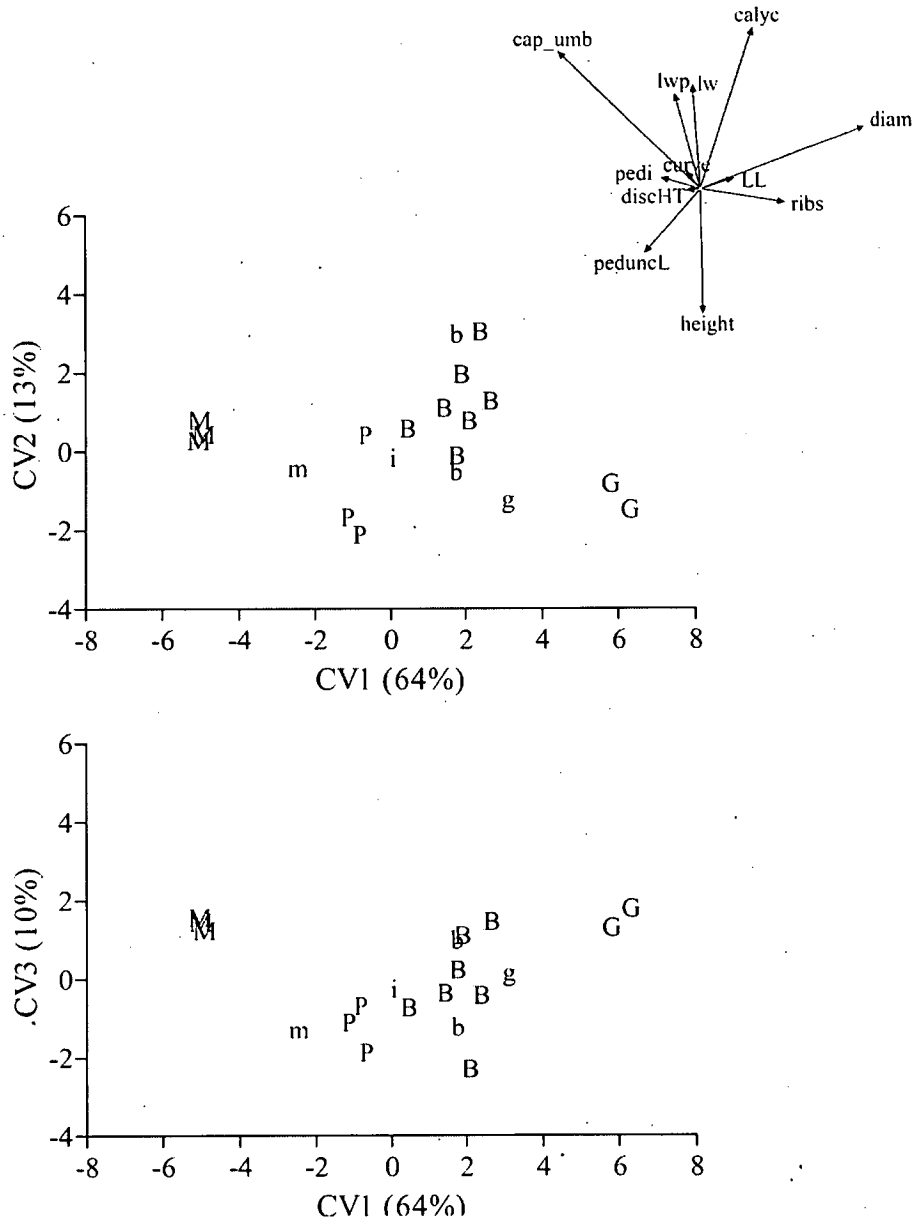


Figure 1.25. Plots of *E. globulus* region mean canonical variate scores, from analysis using 12 capsule, inflorescence and leaf variables. The proportion of variation explained by each axis is shown. The vector diagram in the top corner indicates the direction of the variables in the space defined by CV1 and CV2 and the length of the vector reflects the importance of each variable in maximising differences between groups. B, *bicostata*; P, *pseudoglobulus*; G, *globulus*; M, *maidenii*; m, *maidenii-pseudoglobulus* intergrade; b, *bicostata-pseudoglobulus* intergrade; i, *bicostata-maidenii-pseudoglobulus* intergrade; g, *globulus* intergrades. Some key regions are identified by the numeric codes in Table 1.2.

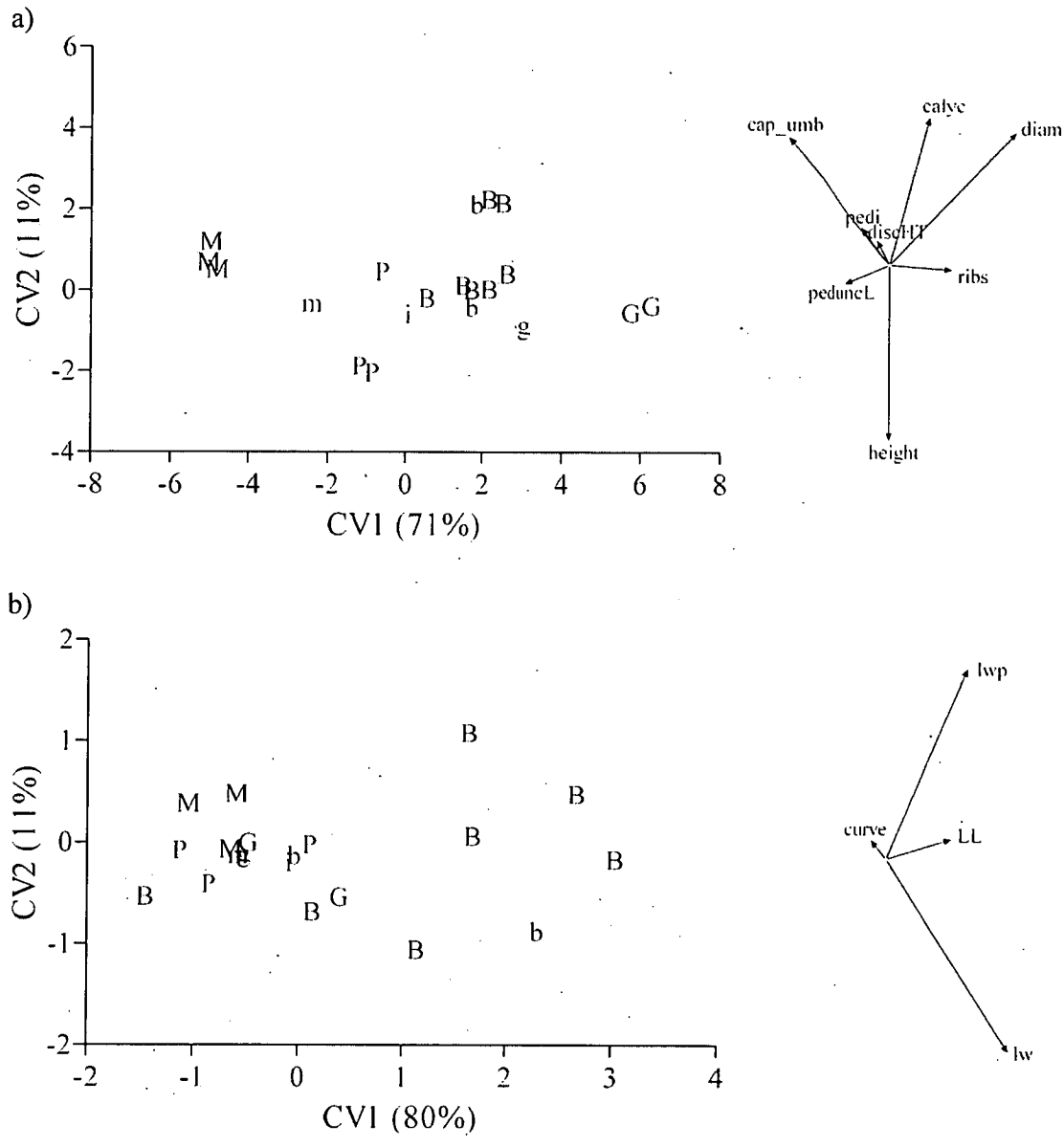


Figure 1.26. Plots of *E. globulus* region mean canonical variate scores, from analysis using a) eight capsule and inflorescence variables and b) four leaf variables. The proportion of variation explained by each axis is shown. The vector diagram indicates the direction of the variables in the space defined by CV1 and CV2 and the length of the vector reflects the importance of each variable in maximising differences between groups. B, *bicostata*; P, *pseudoglobulus*; G, *globulus*; M, *maidenii*; m, *maidenii-pseudoglobulus* intergrade; b, *bicostata-pseudoglobulus* intergrade; i, *bicostata-maidenii-pseudoglobulus* intergrade; g, *globulus* intergrades.

1.3.2.2 Morphological affinities of geographically outlying regions or localities

Geographically outlying regions of *bicostata* at 1-Mt Bryan, 2-Mt Cole, 6-Jenolan and 7-Wollemi formed part of the morphological core of *bicostata* but tended to deviate in the direction of *pseudoglobulus* (Figure 1.25), except 1-Mt Bryan that deviated away from both *bicostata* and *pseudoglobulus* in the third dimension (Figure 1.25). The geographically outlying 14-Lerderderg Gorge region, which is classified as *pseudoglobulus*, deviated away from other core *pseudoglobulus* regions towards *bicostata* (Figure 1.25). The isolated locality at 30-Nadgee River, originally collected as *maidenii-pseudoglobulus* intergrade, fell within the *pseudoglobulus* morphological core (Figure 1.27).

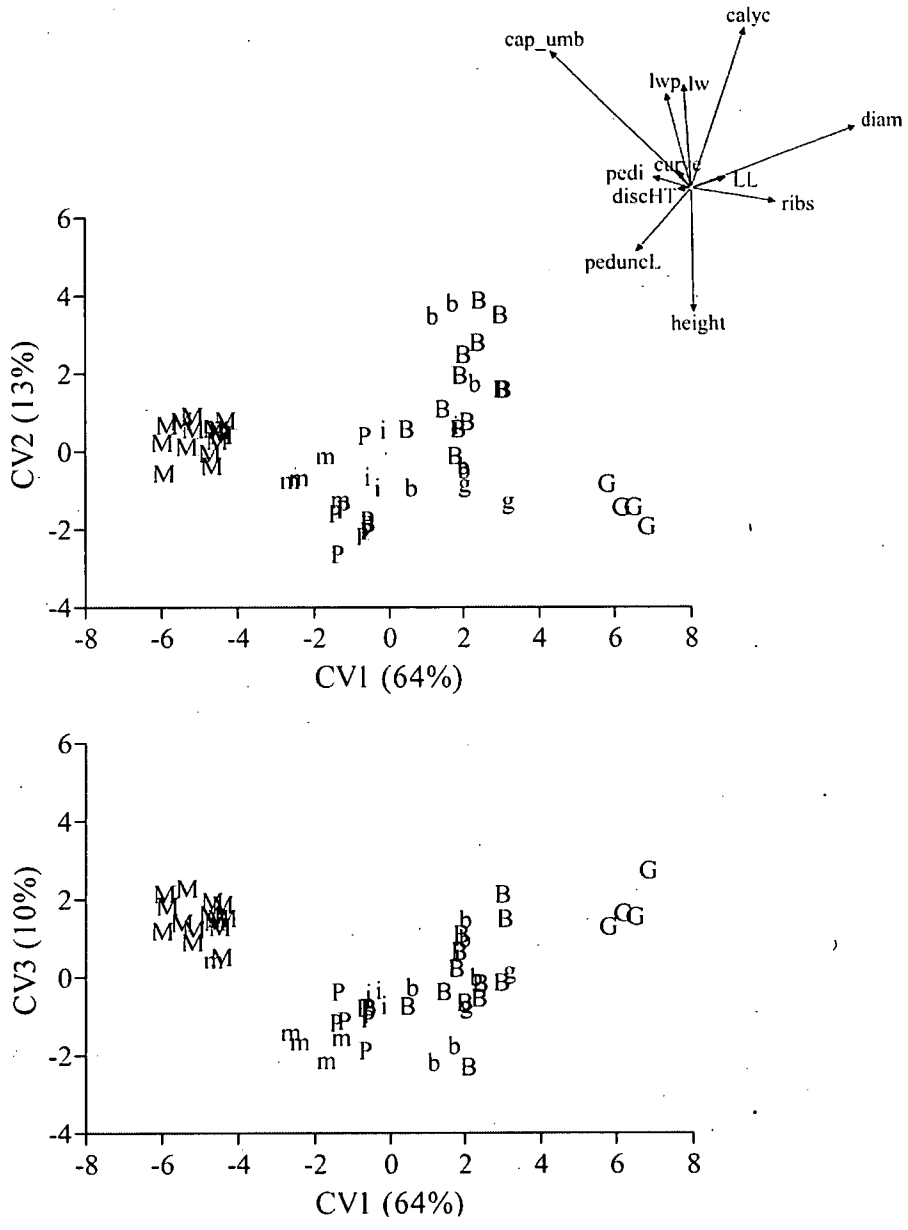


Figure 1.27. Plots of *E. globulus* locality mean canonical variate scores in the space defined by the region level analysis, using 12 capsule, inflorescence and leaf variables. The proportion of variation explained by each axis is shown. The vector diagram indicates the direction of the variables in the space defined by CV1 and CV2 and the length of the vector reflects the importance of each variable in maximising differences between groups. B, *bicostata*; P, *pseudoglobulus*; G, *globulus*; M, *maidenii*; m, *maidenii-pseudoglobulus* intergrade; b, *bicostata-pseudoglobulus* intergrade; i, *bicostata-maidenii-pseudoglobulus* intergrade; g, *globulus* intergrades. Some localities are identified by the numeric codes in Table 1.2.

1.3.2.3 Geographic extent of morphologically core populations

While trees from 16-Omeo region were originally collected as intergrade *bicostata-pseudoglobulus*, the morphometric analysis showed that individuals in this region generally possessed the core *bicostata* morphology (Figure 1.25). Similarly, the locality at 31-Maramingo Creek, collected as *maidenii-pseudoglobulus* intergrade, fell within the *maidenii* core (Figure 1.27). The locality level analysis also showed that the *maidenii* core included the northern part of the Cann Valley highway (locality 28) but the central and southern localities on the highway (localities 33 and 34) were intergrade *maidenii-pseudoglobulus*, as was the locality at 32-Alfred NP (Figure 1.27). The localities collected as core *pseudoglobulus* (apart from 43-Lerderderg) clustered together but this group was not highly differentiated from the rest of the complex, unlike the core localities of *maidenii*, *bicostata* and *pseudoglobulus* that formed the extremities of the variation in the ordination (Figure 1.27). Rather, the *pseudoglobulus* localities had morphological affinities to the intergrade localities that form part of the continuum between *bicostata* and *maidenii* (Figure 1.27).

1.3.2.4 Morphological affinities of intergrade populations

While the key taxa of the *E. globulus* complex were differentiated, a morphological continuum of intergrade populations linked the core taxa. In the 15-Buchan region, southern localities (44-Stoney Creek and 46-Buchan) were intermediate between *pseudoglobulus*, *maidenii* and *bicostata* but with closest morphological affinities to *pseudoglobulus* (Figure 1.27). Northern and western localities in this region (47-Gelantipy and 45-Cutts Creek Rd) had closest morphological affinities to *bicostata* (Figure 1.27). Intergrade localities in Gippsland were morphologically intermediate between *globulus* and the continuum between *pseudoglobulus* and *bicostata* (Figure 1.27). This included localities collected in the Mitchell River (51-Cobbannah and 52-Peel Gap) and Strzelecki Ranges (54-Bowden-Carrajung and 55-Jeeralang) regions (Figure 1.27). The locality at 53-Heyfield in Gippsland was also intermediate but had closer affinities to *pseudoglobulus* (Figure 1.27).

1.3.3 Relationships between morphological and molecular affinities

1.3.3.1 Relationship between morphological, geographic and molecular affinities of regions across the *E. globulus* species complex

There was a highly significant association between Nei's (1978) genetic distance and geographic distance (Figure 1.28). Outliers on this plot were the pairwise comparisons of regions that were separated by a strong geographic barrier over a short geographic distance (e.g. Great Dividing Range). There was no association between Mahalanobis distance and geographic distance (Figure 1.29). Nei's (1978) genetic distance and Mahalanobis distance were weakly but significantly correlated (Figure 1.30). Outliers on this plot could mostly be attributed to the pairwise comparison of pure Victorian *globulus* populations with nearby but morphologically differentiated populations.

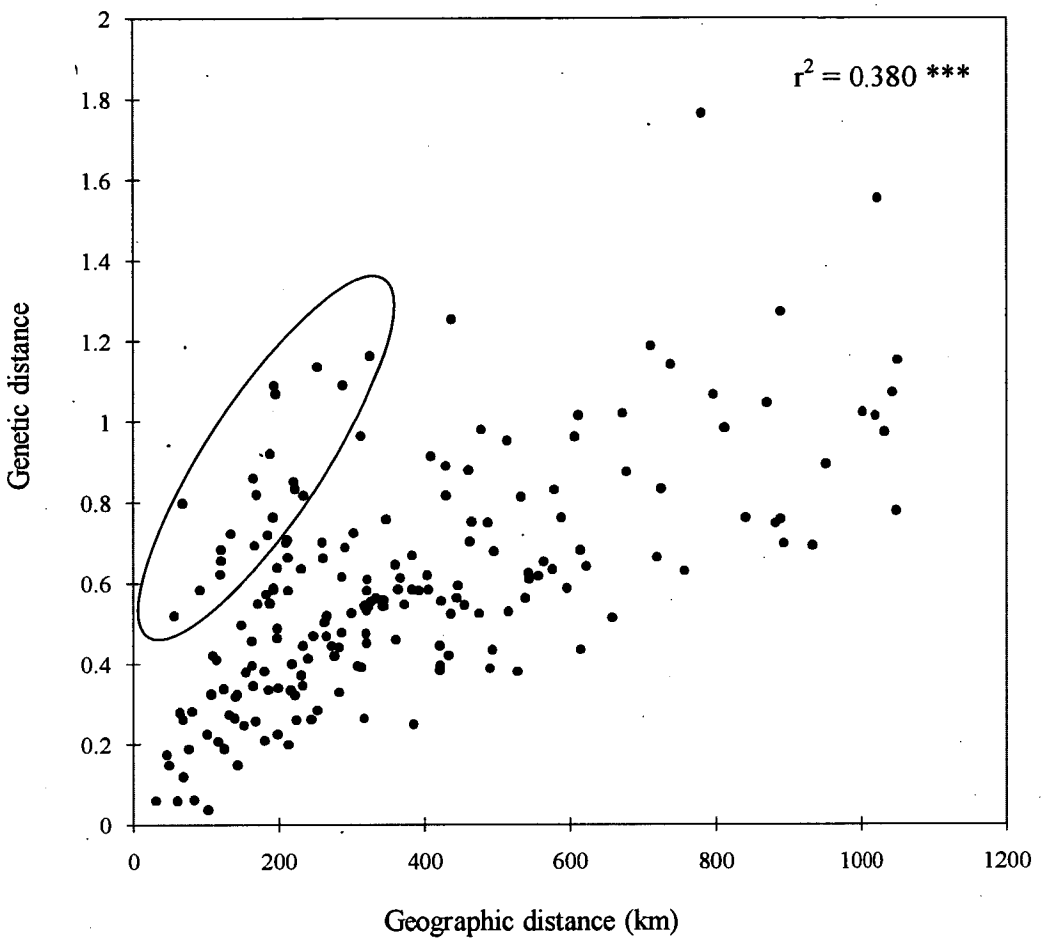


Figure 1.28. Relationship between geographic distance and Nei's (1978) genetic distance between regions of the *E. globulus* complex. Only regions for which both morphological and molecular data were available are shown. The Mantel test and its significance are shown. The ellipse contains those regions that are separated by geographic barriers over a short distance (e.g. Great Dividing Range).

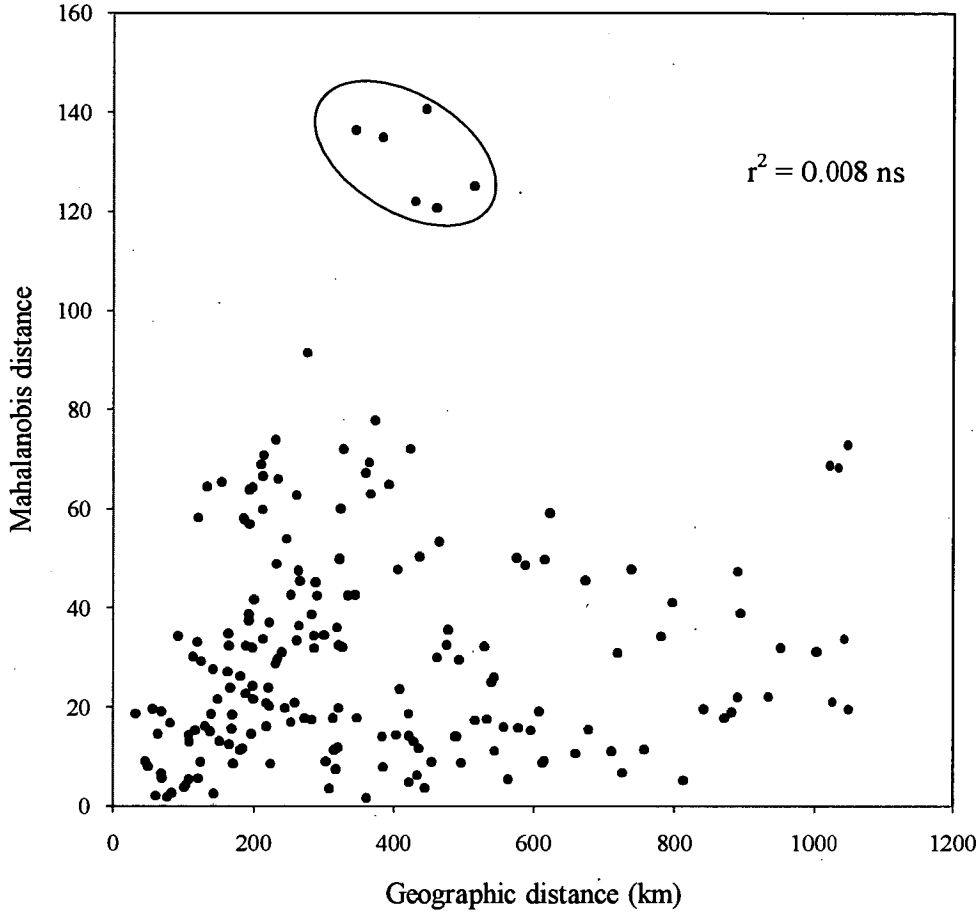


Figure 1.29. Relationship between geographic distance and Mahalanobis distance between regions of the *E. globulus* complex. Only regions for which both morphological and molecular data were available are shown. The Mantel test and its significance are shown. The ellipse contains the pairwise comparisons between Victorian *globulus* and *maidenii*.

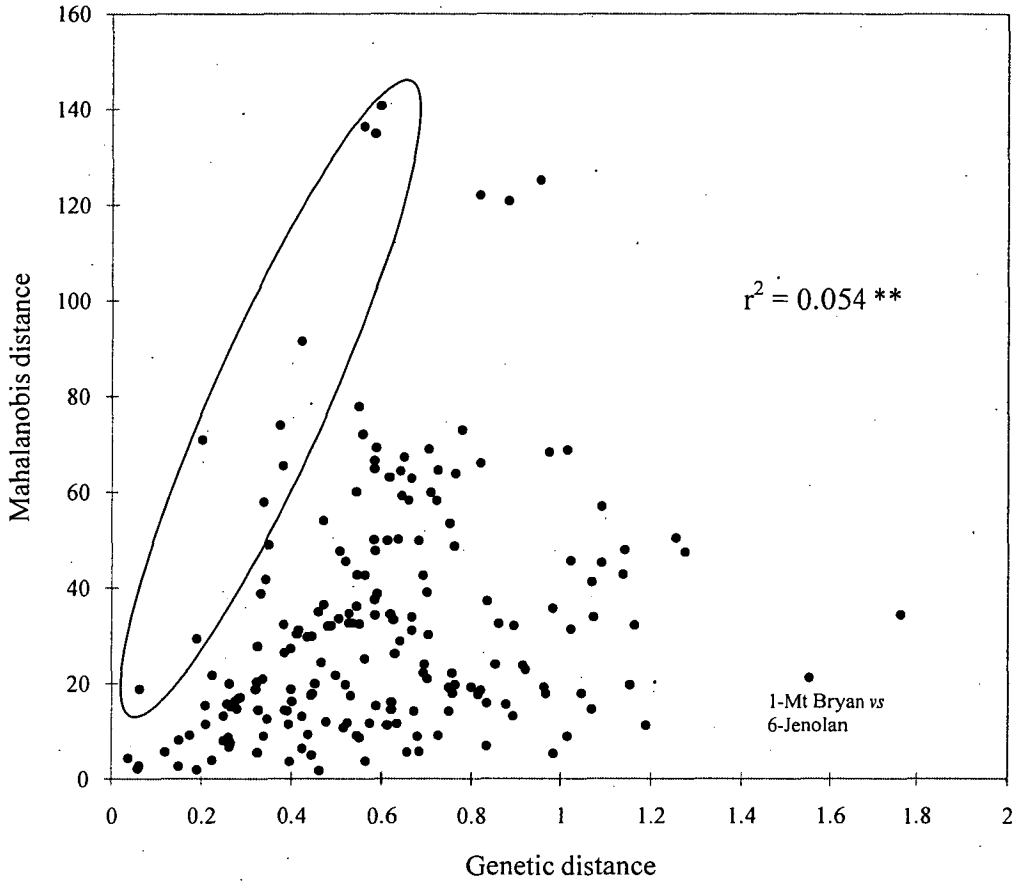


Figure 1.30. Relationship between Nei's (1978) genetic distance and Mahalanobis distance between regions of the *E. globulus* complex. Only regions for which both morphological and molecular data were available are shown. The Mantel test and its significance are shown. The ellipse contains the pairwise comparisons between Victorian *globulus* and *maidenii*, *pseudoglobulus* and the Victorian intergrades.

1.3.3.2 Correlations between morphological and molecular affinities in NSW and Victorian populations of the *E. globulus* species complex

As a recent hybrid zone would be expected to have correlated morphological and molecular affinities, the morphological and molecular affinities of individuals in regions of NSW and Victoria were tested for significant correlations in each region. There were significant associations between the same molecular and morphological affinities in four regions (Table 1.9) and these relationships were graphed (see below).

Table 1.9. Pearson's and Spearman's correlations (r) between affinities of *E. globulus* individuals within NSW/Vic. regions to each of the four taxa based on microsatellites and the affinities of individuals to the equivalent taxon based on morphology. Regions with $n < 10$ are excluded. Only significant associations are shown (*** $P < 0.001$; ** $0.001<P<0.01$; * $0.01<P<0.05$).

Region (taxon)	n	Pearson	Spearman
2 Mt Cole (B)	15		
3 Hume (B)	26		
4 NE Victoria (B)	22		
8 Araluen (M)	25		
9 Wadbilliga (M)	34	0.38*(M)	
10 SE Forests (M)	20		
11 Alfred-Nadgee (m-p)	42	0.51***(P); 0.31*(M)	0.53***(P); 0.37*(M)
12 Mt Cann (P)	22		
13 Lakes Entrance (P)	44		
14 Lerderderg Gorge (P)	16		
15 Buchan (b-m-p)	33		
16 Omeo (b-p)	21		
17 Mitchell River (b-p)	21		
18 Strzelecki Ranges (g)	25		0.57** (G)
24 South Gippsland (G)	26		-0.44*(G); -0.43*(B)

1.3.3.2.1 Alfred-Nadgee

In the 11-Alfred-Nadgee region, there was a significant correlation between *pseudoglobulus*-type morphology and *pseudoglobulus*-type microsatellite genotype, and there was also a correlation between *maidenii*-type morphology and *maidenii*-type microsatellite genotype (Table 1.9). This correlation was due to a subset of individuals in this region having strong morphological affinities to *pseudoglobulus* and these individuals also having a molecular tendency towards *pseudoglobulus*, and a different subset of individuals having morphological affinities to *maidenii* and molecular tendencies towards this taxon (Figure 1.31) as is the pattern in core populations of these two taxa (Figure 1.32), rather than there being a large group of individuals with morphological and molecular affinities intermediate between *pseudoglobulus* and *maidenii* (Figure 1.31).

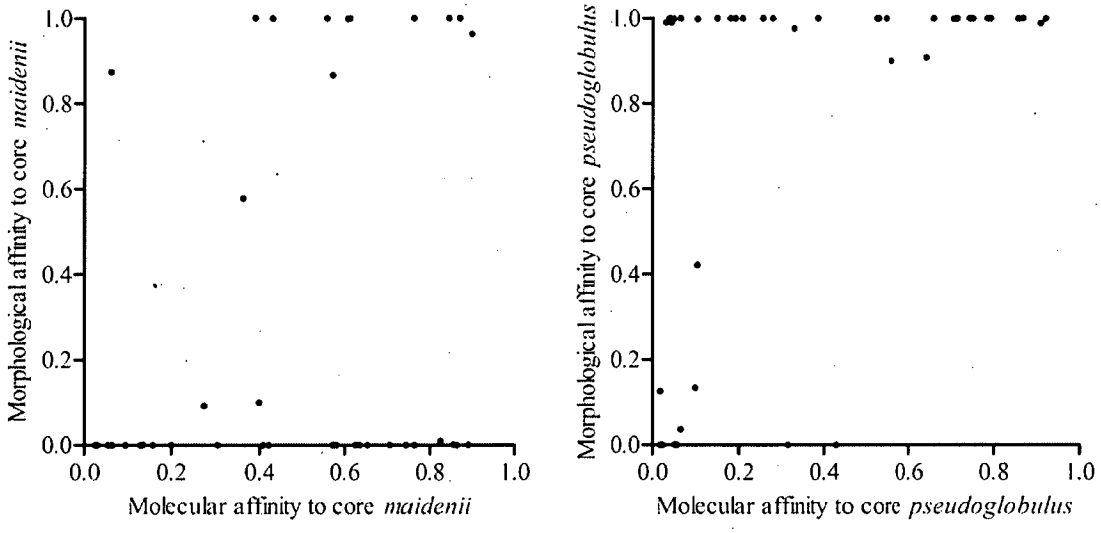


Figure 1.31. Relationship between molecular (Q) and morphological affinities of individuals to *pseudoglobulus* and *maidenii* in the 11-Alfred-Nadgee m-p intergrade region. STRUCTURE analyses for calculation of Q were performed using K=4 and *a priori* population groupings corresponding to the four core taxa, and an admixture model for remaining individuals.

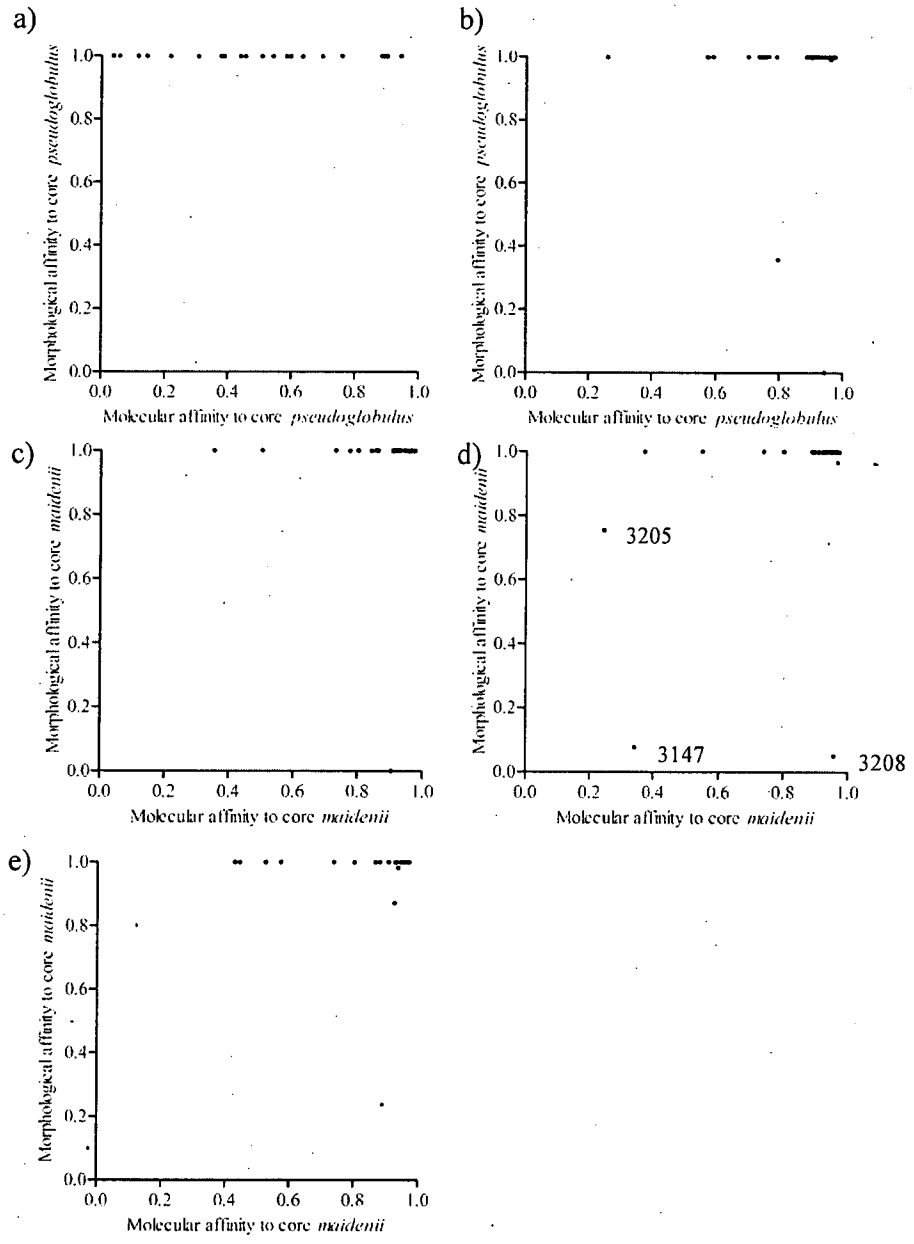
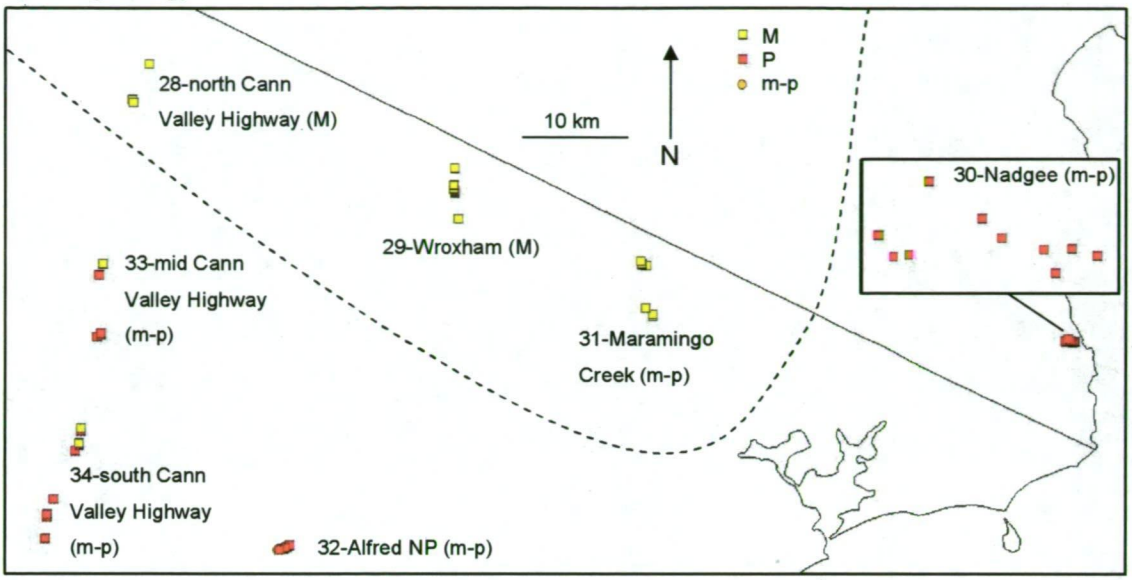


Figure 1.32. Relationship between molecular (Q) and morphological affinities of individuals to *pseudoglobulus* in the core *pseudoglobulus* regions a) 12-Mt Cann; b) 13-Lakes Entrance and the relationship between molecular (Q) and morphological affinities of individuals to *maidenii* in the core *maidenii* regions c) 8-Araluen; d) 9-Wadbilliga; e) 10-SE Forests. Tree numbers for outliers in the 9-Wadbilliga region are shown. Tree 3147 is from the 23-Brown Mountain locality; 3205 and 3208 are from 22-Mumbulla. STRUCTURE analyses for calculation of Q were performed using K=4 and *a priori* population groupings corresponding to the four core taxa, and an admixture model for remaining individuals.

In the 11-Alfred-Nadgee region, there was geographic clustering of individuals by morphological and molecular affinities. The *pseudoglobulus*-type individuals (in terms of both microsatellites and morphology) were from the 30-Nadgee locality within this region, and *maidenii*-type individuals at the 31-Maramingo Creek locality (Figure 1.33), though both of these localities were classified as m-p morphology in the field. At the northern and southern ends of the Cann Valley Highway, there was an association between morphology and molecular affinity, with *maidenii*-like individuals in the north and *pseudoglobulus*-type individuals in the south (Figure 1.33). However, there was no association between morphology and molecular affinity in the individuals that linked the northern and southern ends of the highway (Figure 1.33). Individuals at the 32-Alfred NP locality had *pseudoglobulus* morphological affinities but mixed molecular affinities (Figure 1.33).

11-Alfred-Nadgee

a) Morphology



b) Microsatellites

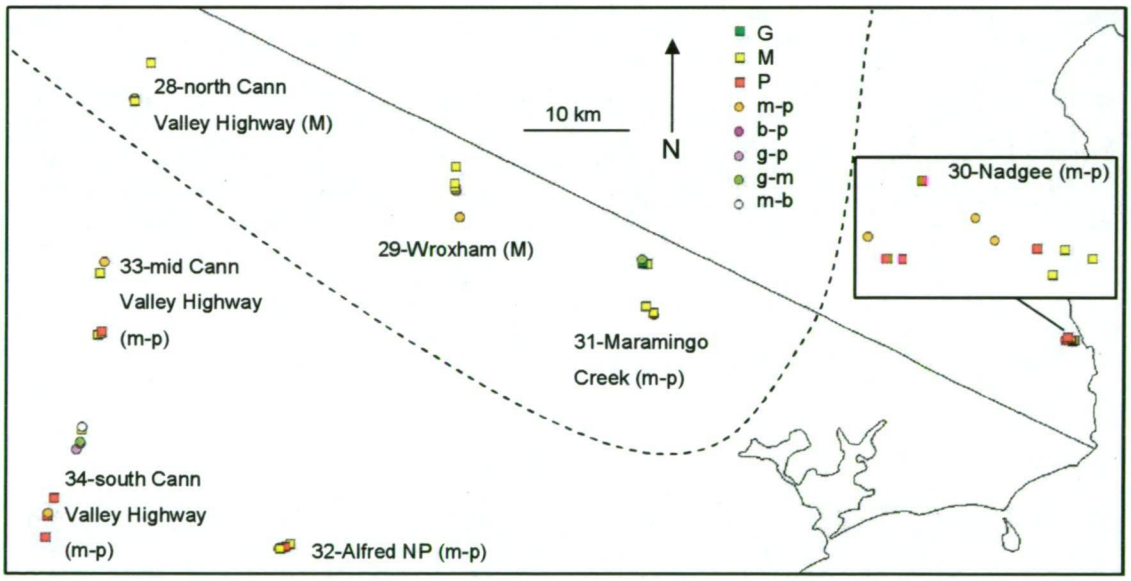


Figure 1.33. Map of the a) morphological and b) molecular affinities of individuals in the 11-Alfred-Nadgee m-p intergrade region (including some pure *maidenii* populations). This map is an inset of the map on p.16. STRUCTURE analyses were performed using K=4 and *a priori* population groupings corresponding to the four core taxa, and an admixture model for remaining individuals. See 1.2.4.2 for the criteria for allocation to the morphological and molecular groups. Each symbol is a single individual, and only individuals for which both morphological and molecular data were available are shown. The southern extent of the *maidenii* gene pool is indicated by a dashed line.

1.3.3.2.2 Wadbilliga

There was a significant correlation between *maidenii*-type morphology and *maidenii*-type microsatellite genotype at the core *maidenii* region 9-Wadbilliga (Table 1.9), though this was only significant using Pearson's correlation. The three outlying individuals were from the 22-Mumbulla and 23-Brown Mountain localities (Figure 1.32) and they had morphological and molecular affinities towards *pseudoglobulus* instead of to *maidenii*, with fruit in groups of three instead of seven (data not shown).

1.3.3.2.3 Strzelecki Ranges and South Gippsland

There was a strong correlation between morphological and molecular affinity to *globulus* in the 18-Strzelecki Ranges region (Table 1.9). Seven individuals had strong and similar morphological and molecular affinities to *globulus* in this region (Figure 1.34), similar to the core population of this taxon (Figure 1.35). However, these individuals were not geographically clustered (Figure 1.36) and there were also more individuals with molecular affinities to *globulus* but no morphological affinities to this taxon (Figure 1.34).

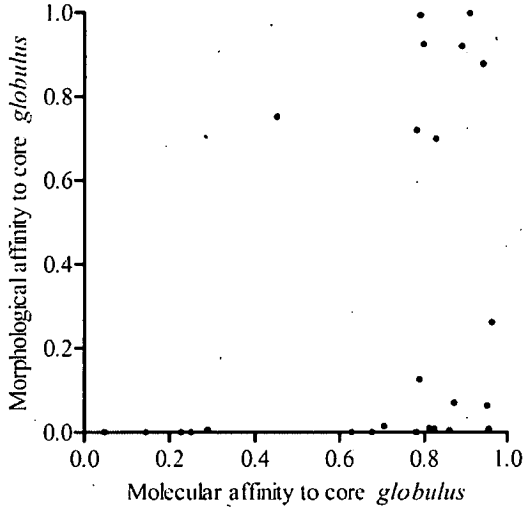


Figure 1.34. Relationship between molecular (Q) and morphological affinities of individuals to *globulus* in the 18-Strzelecki Ranges region. STRUCTURE analyses for calculation of Q were performed using K=4 and *a priori* population groupings corresponding to the four core taxa, and an admixture model for remaining individuals.

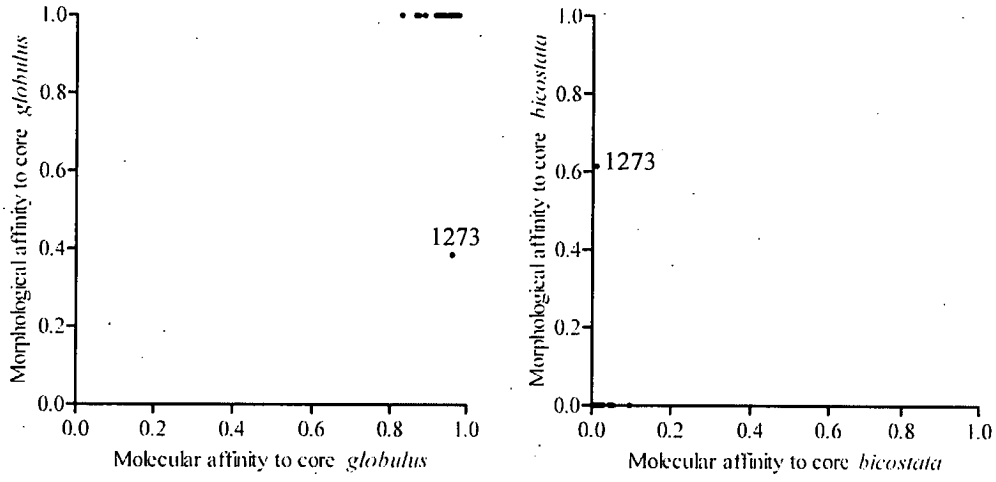
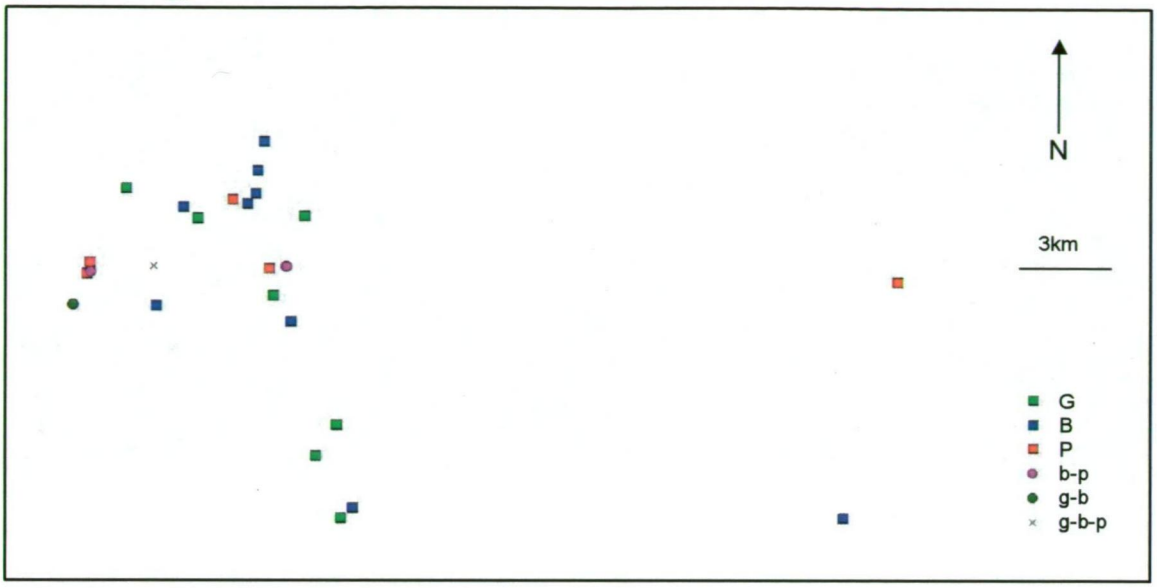


Figure 1.35. Relationship between molecular (Q) and morphological affinities of individuals to *globulus* and *bicostata* in the core Victorian *globulus* region 24-South Gippsland. STRUCTURE analyses for calculation of Q were performed using K=4 and *a priori* population groupings corresponding to the four core taxa, and an admixture model for remaining individuals. The tree number of the outlying individual is shown. This individual is from the 71-Alberton West locality.

18-Strzelecki

a) Morphology



b) Microsatellites

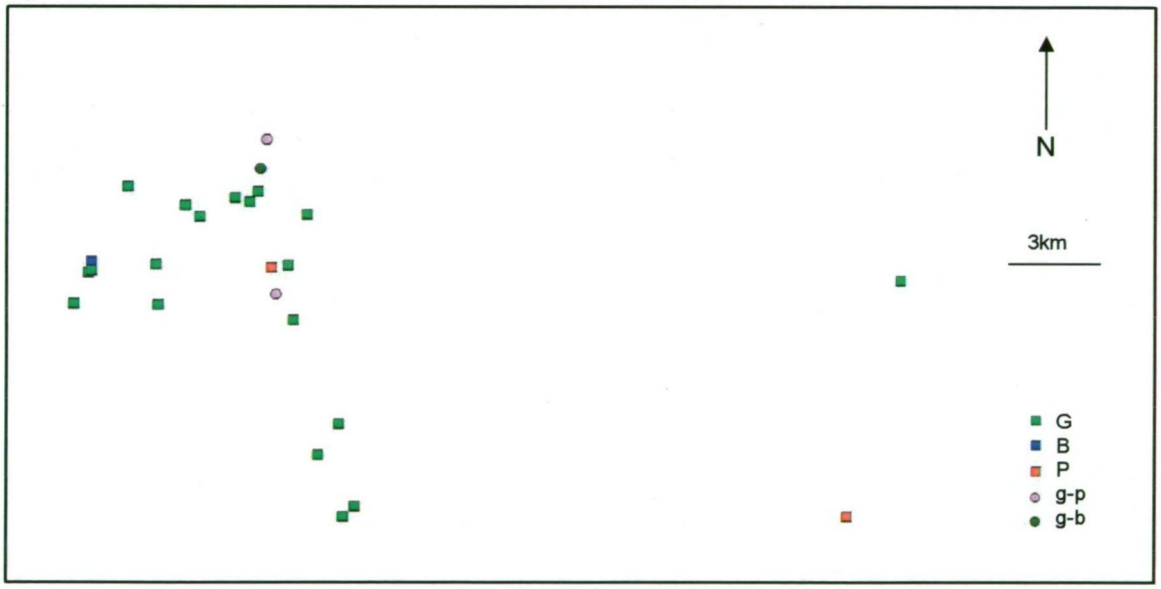


Figure 1.36. Map of the a) morphological and b) molecular affinities of individuals in the 18-Strzelecki Ranges intergrade region. This map is an inset of the map on p.16. STRUCTURE analyses for calculation of Q were performed using K=4 and *a priori* population groupings corresponding to the four core taxa, and an admixture model for remaining individuals. See 1.2.4.2 for the criteria for allocation to the morphological and molecular groups. Each symbol is a single individual, and only individuals for which both morphological and molecular data were available are shown.

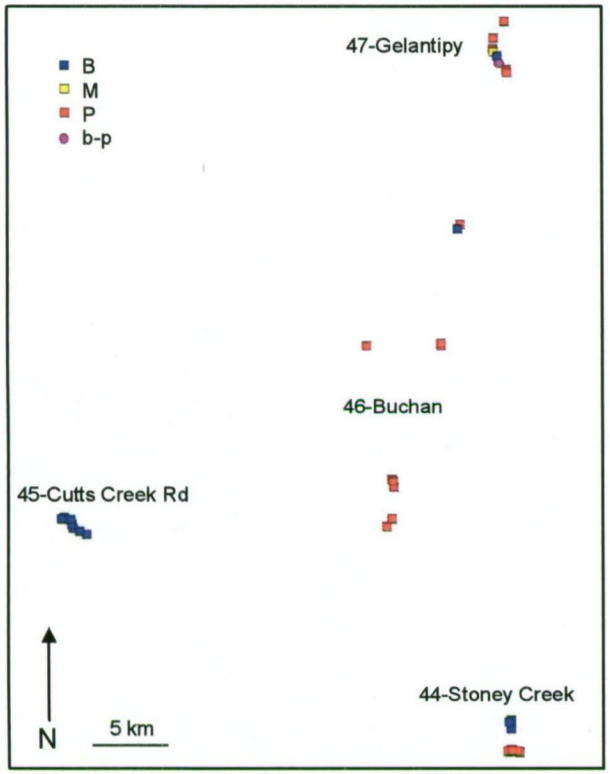
There was a significant association between *globulus*-type morphology and *globulus*-type microsatellite genotype in the core *globulus* 24-South Gippsland region, as well as an association between *bicostata*-type morphology and *bicostata*-type microsatellite genotype (Table 1.9). However this was driven by a single individual (1273) from the 71-Alberton West locality (Figure 1.35) that had smaller but solitary fruit (data not shown).

1.3.3.2.4 Buchan

At the 15-Buchan region, there was no significant correlation between morphological and molecular affinities (Table 1.9). There was geographical clustering of morphotypes in this region; individuals with *bicostata* morphological affinities were clustered in the west (locality 45-Cutts Creek) and *pseudoglobulus*-like individuals were located in eastern and northern populations (Figure 1.37). Individuals with molecular affinities to *bicostata* and *pseudoglobulus* were, however, spread geographically throughout the region (Figure 1.37).

15-Buchan

a) Morphology



b) Microsatellites

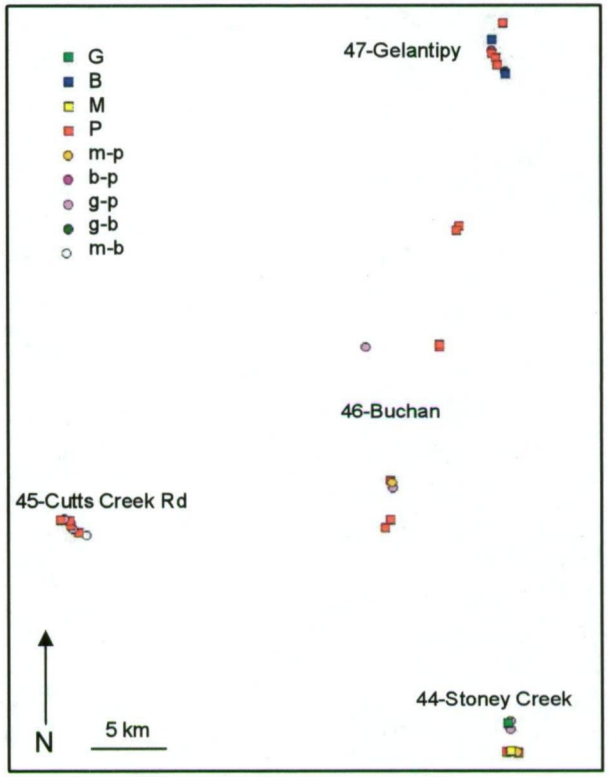


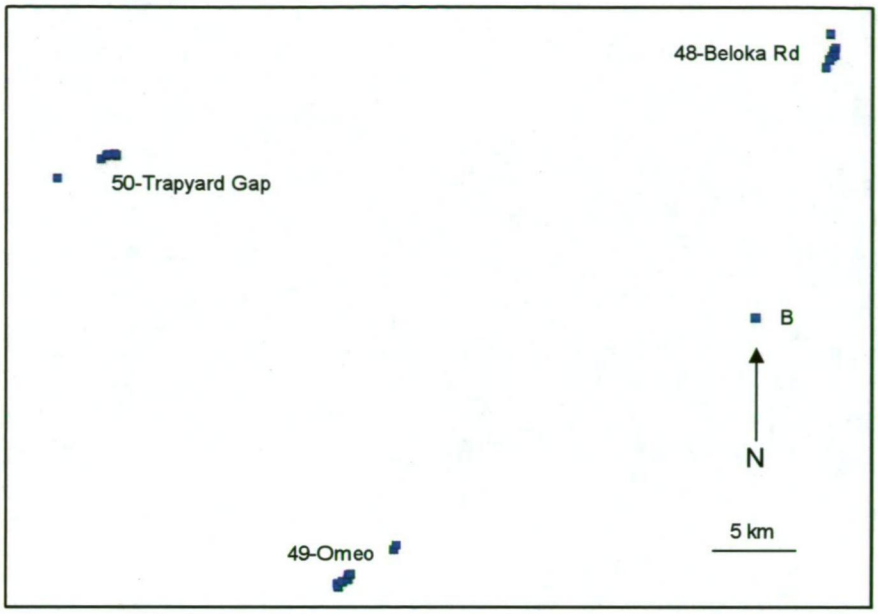
Figure 1.37. Map of the a) morphological and b) molecular affinities of individuals in the 15-Buchan b-m-p intergrade region. This map is an inset of the map on p.16. STRUCTURE analyses for calculation of Q were performed using K=4 and *a priori* population groupings corresponding to the four core taxa, and an admixture model for remaining individuals. See 1.2.4.2 for the criteria for allocation to the morphological and molecular groups. Each symbol is a single individual, and only individuals for which both morphological and molecular data were available are shown.

1.3.3.2.5 Omeo and Mitchell River (b-p intergrades)

The two b-p intergrade regions (16-Omeo and 17-Mitchell River) had contrasting morphological and molecular affinities. Individuals in the 16-Omeo region had strong morphological and molecular affinities to *bicostata* (Figure 1.38), similar to core populations of this taxon, so this region is essentially part of the *bicostata* gene pool and was either misclassified or taxonomic characters were weighted differently in the field. On the other hand, individuals of the 17-Mitchell River region had morphological affinities mostly to *bicostata* and *pseudoglobulus*, with microsatellite affinities mostly to *globulus* (Figure 1.39). There was no significant association between morphology and molecular type in the 16-Omeo or 17-Mitchell River regions (Table 1.9).

16-Omeo

a) Morphology



b) Microsatellites

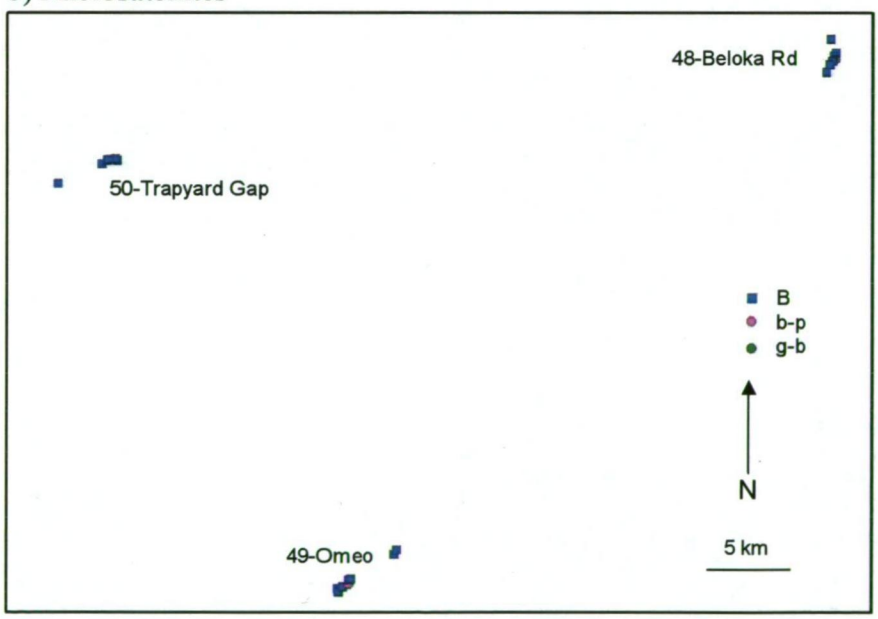
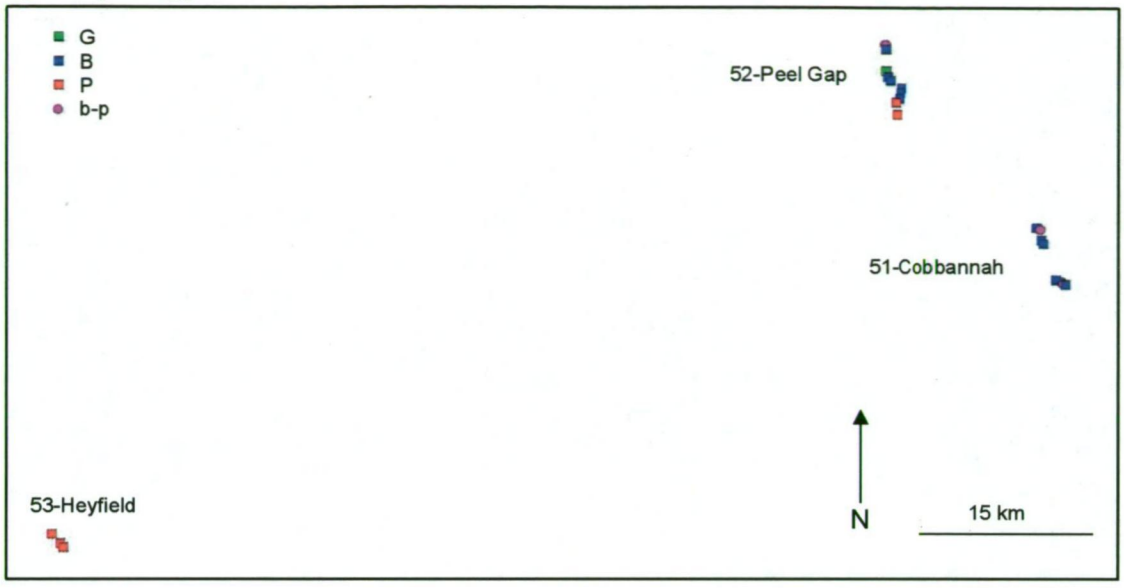


Figure 1.38. Map of the a) morphological and b) molecular affinities of individuals in the 16-Omeo b-p intergrade region. This map is an inset of the map on p.16. STRUCTURE analyses for calculation of Q were performed using K=4 and *a priori* population groupings corresponding to the four core taxa, and an admixture model for remaining individuals. See 1.2.4.2 for the criteria for allocation to the morphological and molecular groups. Each symbol is a single individual, and only individuals for which both morphological and molecular data were available are shown.

17-Mitchell River

a) Morphology



b) Microsatellites

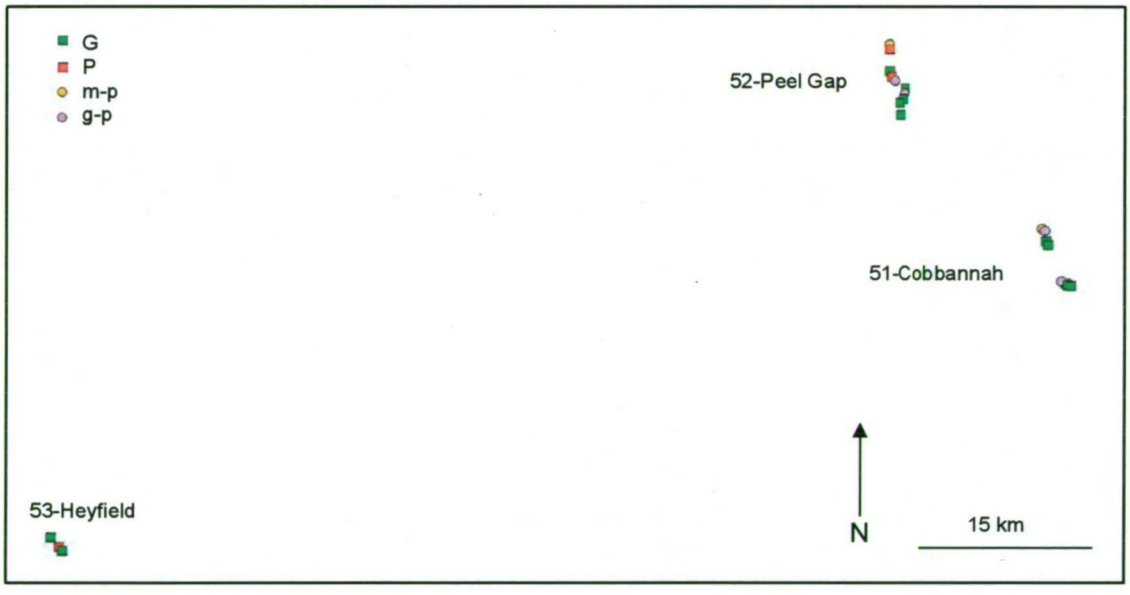


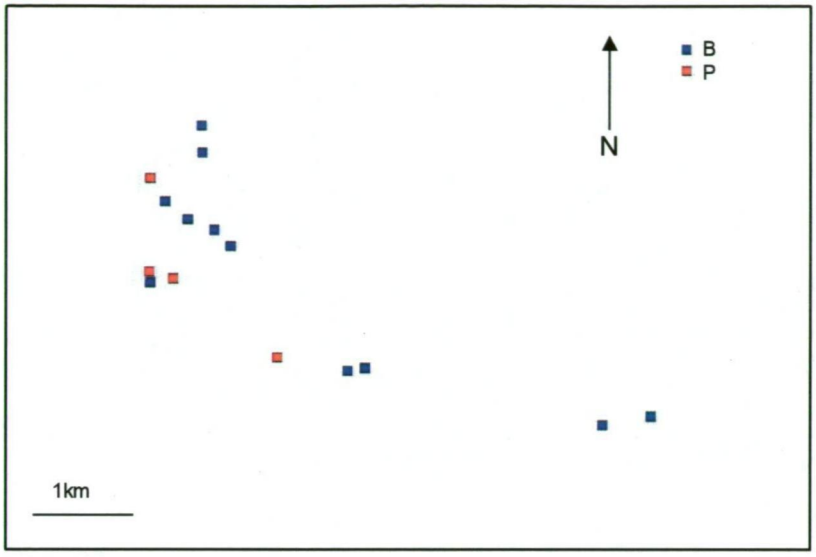
Figure 1.39. Map of the a) morphological and b) molecular affinities of individuals in the 17-Mitchell River b-p intergrade region. This map is an inset of the map on p.16. STRUCTURE analyses for calculation of Q were performed using K=4 and *a priori* population groupings corresponding to the four core taxa, and an admixture model for remaining individuals. See 1.2.4.2 for the criteria for allocation to the morphological and molecular groups. Each symbol is a single individual, and only individuals for which both morphological and molecular data were available are shown.

1.3.3.2.6 Mt Cole

Individuals at 2-Mt Cole had strong morphological affinities to *bicostata*, with a small number of morphological outliers (Figure 1.40, Figure 1.41). At the molecular level this population was admixed between *bicostata* and *pseudoglobulus* (Figure 1.40), which is in contrast to the pattern observed in core *bicostata* (data not shown). There was no significant association between morphology and molecular type in this region (Table 1.9).

2-Mt Cole

a) Morphology



b) Microsatellites

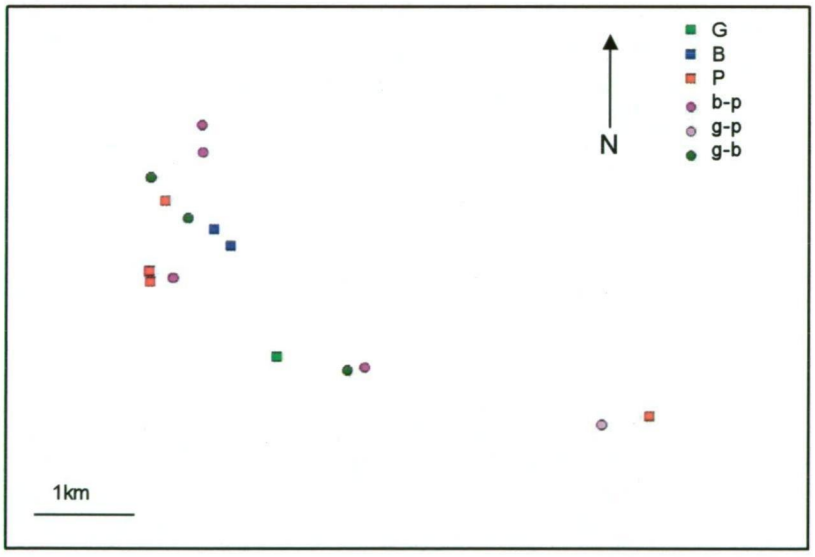


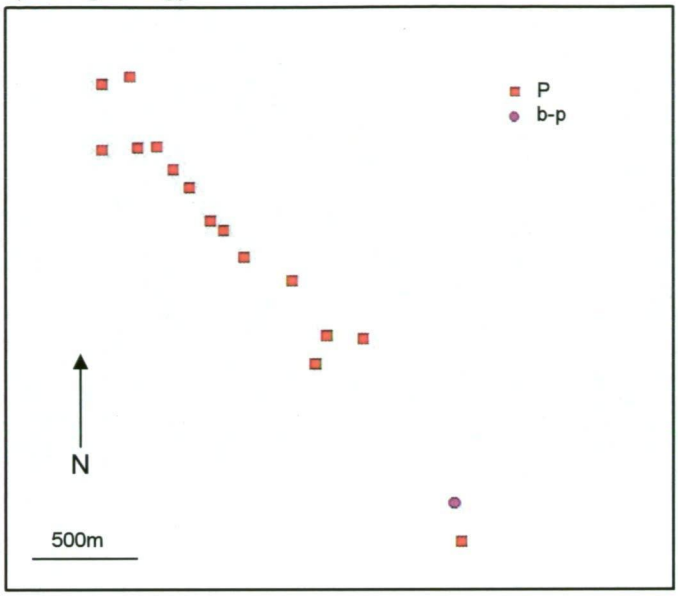
Figure 1.40. Map of the a) morphological and b) molecular affinities of individuals in the 2-Mt Cole region. This map is an inset of the map on p.16. STRUCTURE analyses for calculation of Q were performed using K=4 and *a priori* population groupings corresponding to the four core taxa, and an admixture model for remaining individuals. See 1.2.4.2 for the criteria for allocation to the morphological and molecular groups. Each symbol is a single individual, and only individuals for which both morphological and molecular data were available are shown.

1.3.3.2.7 Lerderderg Gorge

Individuals at 14-Lerderderg Gorge had strong morphological affinities to *pseudoglobulus*, with only one morphological outlier (Figure 1.40, Figure 1.41). At the molecular level this populations was admixed between *globulus*, *bicostata* and *pseudoglobulus* (Figure 1.41), in contrast to the affinities of populations of core *pseudoglobulus* (Figure 1.32). There was no significant association between morphology and molecular type in this region (Table 1.9).

14-Lerderderg

a) Morphology



b) Microsatellites

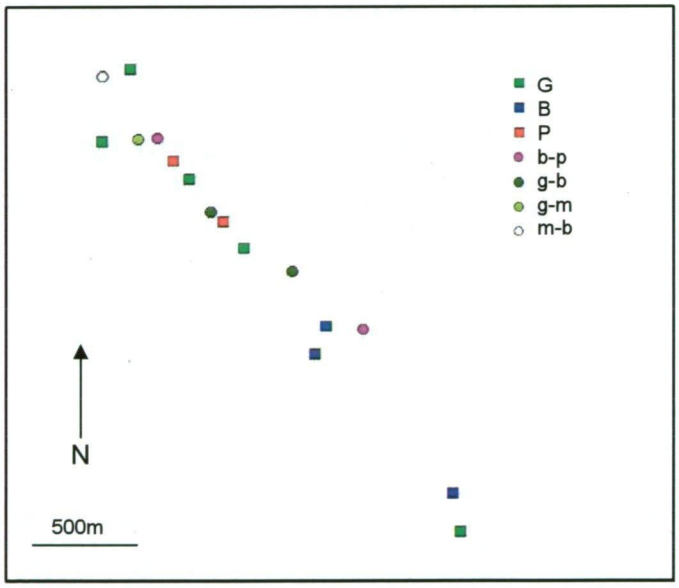


Figure 1.41. Map of the a) morphological and b) molecular affinities of individuals in the Lerderderg region. This map is an inset of the map on p.16. STRUCTURE analyses for calculation of Q were performed using K=4 and *a priori* population groupings corresponding to the four core taxa, and an admixture model for remaining individuals. See 1.2.4.2 for the criteria for allocation to the morphological and molecular groups. Each symbol is a single individual, and only individuals for which both morphological and molecular data were available are shown.

1.4 Discussion

1.4.1 Patterns of diversity in the E. globulus *complex compared to other eucalypts*

The high levels of diversity estimated here for the whole *E. globulus* complex (H_e mean across populations = 0.78, H_e population range = 0.57-0.87, H_T = 0.89) were comparable to genetic diversity measures in other microsatellite studies of natural populations of eucalypts, such as *E. camaldulensis* (H_e mean and range = 0.83, 0.67-0.89, Butcher *et al.* 2009), E. *urophylla* in Indonesia (H_e mean and range = 0.74, 0.70-0.78, Payn *et al.* 2008), *E. wandoo* in Western Australia (H_T = 0.87, Byrne *et al.* 2008), *E. brownii-populnea* in Queensland (H_e mean and range = 0.88, 0.84-0.91, Holman *et al.* 2003), *E. benthamii* in NSW (H_e mean and range = 0.70, 0.618-0.747, Butcher *et al.* 2005), *E. perriniana* (H_e mean and range = 0.73, 0.65-0.83, H_T = 0.85, Rathbone *et al.* 2007) and *E. vernicosa-subcrenulata-johnstonii* in Tasmania (H_T = 0.86, McGowen *et al.* 2001). The lower diversity levels in populations of *E. curtisii* (H_e mean and range = 0.54, 0.30-0.67, Smith *et al.* 2003) and *E. morrisbyi* (H_e mean and range = 0.69, 0.67-0.72 Jones *et al.* 2005) can be explained by the localised distribution of these two eucalypt species, a pattern also reported in isozyme studies (Potts and Wiltshire 1997).

As expected, Wright's Fixation Index was positive in nearly all localities and regions. This measure is commonly positive in *Eucalyptus* (Potts and Wiltshire 1997) due to the effects of inbreeding, the Wahlund effect or null alleles. In this study, the positive F values could be due to all three of these factors, as eucalypts have a mixed mating system, with a mean outcrossing rate of 0.74 measured over 23 species (Byrne 2008), fine-scale population structure has been detected within populations of *E. globulus* (Jones *et al.* 2007, Foster *et al.* 2007) and null alleles are relatively common at microsatellite loci (reviewed in Dakin and Avise 2004). For EMBRA19 and EMCRC10, F_{IS} was higher than in other loci, both in localities and in regions, indicating the presence of null alleles at these two loci. The presence of null alleles at EMCRC10 in nearly all regions, including in Tasmania, is somewhat surprising, given that this primer pair was designed from a *globulus* sequence (Steane *et al.* 2001).

Homozygote excess across loci in a single population could be attributed to the Wahlund effect and/or inbreeding, rather than null alleles. In this study, some regions (9-Wadbilliga, 11-Alfred-Nadgee, 30-SE Tasmania and 32 NE Tasmania) had evidence for a null allele (as measured by homozygote excess) across many loci, which could be due to the Wahlund effect in these regions, as they cover a large geographic area compared with many other regions. Indeed, when localities within these regions were tested for homozygote excess, fewer loci showed evidence of homozygote excess than in the regional level analysis.

While null alleles can severely bias parentage analysis (Pemberton *et al.* 1995), they may cause less bias in studies of population structure (Dakin and Avise 2004), such as in this study. However, many authors still caution against the use of loci with null alleles in population studies (Chapuis and Estoup 2007), because null alleles can reduce the within-population genetic diversity (Paetkau and Strobeck 1995) and therefore overestimate F_{ST} and genetic distance, since these measures increase as within-population diversity decreases (Slatkin 1995, Hedrick 1999, Jost 2008). In future population genetics or parentage studies in *E. globulus* it may be worth excluding the loci which possibly have null alleles at high frequencies (EMCRC10, EMBRA19).

The moderate F_{ST} of 0.12 (measured at the locality level for the entire *E. globulus* complex) or 0.10 (measured at the regional level) was similar to the G_{ST} of 0.14, averaged across eleven isozyme studies of other widespread eucalypt species (Potts and Wiltshire 1997), but higher than the mean F_{ST} for widespread eucalypts assayed using RFLP and microsatellites (0.062 and 0.055 respectively, Byrne 2008). The overall F_{ST} for *E. globulus* was also higher than that measured in *E. urophylla* using microsatellites (F_{ST} = 0.04, Tripiana *et al.* 2007; F_{ST} = 0.03 Payn *et al.* 2008), despite the disjunct distribution of *E. urophylla* across several islands of Indonesia, and also higher than in *E. camaldulensis*, the most widely distributed eucalypt (F_{ST} = 0.082, Butcher *et al.* 2009). However, as expected, the F_{ST} was lower than in other eucalypts with regional, disjunct distributions, such as *E. perriniana* (F_{ST} = 0.16, Rathbone *et al.* 2007) and *E. curtisii* (F_{ST} = 0.30, Smith *et al.* 2003). The overall F_{ST} in *E. globulus* was higher than the means for widespread eucalypts assayed using RFLP and microsatellites reported in Byrne (2008). This could be due to the large disjunctions in parts of the *E. globulus*

range and the fact that four taxa were included in this study. Indeed, the AMOVA revealed 7% of the variation in the complex was partitioned among the four taxa. Pairwise F_{ST} values among taxa are not commonly cited in the literature, but the pairwise F_{ST} values among taxa (F_{ST} = 0.049-0.099) in this study were mostly higher than the F_{ST} values reported among taxa in other plant species complexes such as oaks (e.g. Aldrich *et al.* 2003, F_{ST} among taxa = 0.01) and *Metrosideros* (Aradhya *et al.* 1991, F_{ST} among taxa = 0.047), which indicates that the taxa of the *E. globulus* species complex have been separated for longer or that barriers to gene flow are stronger than in these other species complexes. In *E. camaldulensis*, the most widespread eucalypt with a distribution covering all states of mainland Australia, an AMOVA revealed that 7.4% of the variation was due to differences among subspecies (Butcher *et al.* 2009). This value is similar to that observed in this study, though the *E. camaldulensis* subspecies are separated by much larger geographic distances than the taxa of the *E. globulus* complex.

Jost (2008) theorised that G_{ST} and its relatives (such as F_{ST}) are defective measures of differentiation as they are affected by levels of genetic diversity within populations and loci, and proposed a new measure of differentiation (D), to allow better comparisons among studies, as it was claimed that traditional measures of differentiation do not always rank species correctly in terms of differentiation. If there are discrepancies between G_{ST}-like statistics and actual levels of differentiation, this could have serious consequences: these values are commonly reported and compared among studies, and are used to draw conclusions about levels of gene flow and relationships between populations, and also to assign conservation priority. While some argue that G_{ST} is still useful in many cases (Ryman and Leimar 2009), others have shown that the use of G_{ST}-like statistics has led to serious misinterpretations in the recent literature (Heller and Siegismund 2009). Heller and Siegismund (2009) calculated D_{est} (Jost 2008) and G'_{ST} (Hedrick 2005) for papers in *Molecular Ecology* that have cited G_{ST}-like statistics [as well as citing k (number of populations) and H_S] over the last two years and showed that G_{ST} underestimated differentiation and would rank two species discordantly around 25% of the time. However, in this study, the rank order of taxa of the *E. globulus* species complex, in terms of differentiation levels, was the same for both F_{ST} and D_{est}. These statistics, calculated for Tasmanian samples of *E. obliqua* (also using microsatellite data)

also rank the same compared to the *E. globulus* taxa whether F_{ST} or D_{est} is used (F_{ST} = 0.016, Bloomfield 2008; D_{est} = 0.188, J. Bloomfield pers. comm.). This is probably because populations of *E. globulus* and *E. obliqua* have similarly high levels of genetic diversity (this study, Bloomfield 2008), and, at least within this *E. globulus* study, the same loci were used across taxa.

The F_{ST} / D_{est} calculated within taxa revealed contrasting patterns of genetic structure among taxa of the *E. globulus* species complex, at times consistent with the distribution of the taxon. The highest differentiation was measured in *bicostata*, which is consistent with its widespread but disjunct distribution. However *globulus*, which is also relatively widespread but disjunct, had less diversity distributed among localities/regions than *bicostata*, indicating the populations of *bicostata* have probably been separated for much longer than those of *globulus*. The low differentiation measured for *maidenii* is somewhat surprising but can be attributed to its continuous distribution and consequently high levels of gene flow among populations. The distribution of *pseudoglobulus* is narrow in the main part of its range in East Gippsland, but includes a disjunct stand 300 km west at Lerderderg Gorge (region 14). However the Lerderderg stand only appears disjunct from *pseudoglobulus* by means of its morphology; these stands have probably been linked by gene flow in the past through other intergrade populations in Victoria, which would account for the low differentiation measured among populations of *pseudoglobulus*.

The central-peripheral hypothesis, that populations at the margins of a species' range exhibit lower genetic diversity and greater genetic differentiation than in the central part of the range (reviewed in Eckert *et al.* 2008), appears to hold at least in *bicostata*, with highest diversity in the continuous, central part of the taxon's range and lower diversity at the edge and in disjunct western (1-Mt. Bryan) and northern (7-Wollemi, and, to a lesser extent, 6-Jenolan) regions. For the species complex as a whole, high levels of genetic diversity were detected in the SE Forests region, indicating that this region is potentially the origin of the species complex. This pattern of decline in genetic diversity with increased distance from the SE Forests region was not strong when compared with, for example, the correlation between SNP haplotype diversity in humans and distance from sub-Saharan Africa (r = -0.91, Li *et al.* 2008). A strong relationship between

latitude and mean number of alleles per allozyme locus was also shown in populations of *Eucalyptus obtusiflora* (r = 0.91, Kennington and James 1998), a widespread species with a linear distribution pattern around 500 km by 100 km. In *E. camaldulensis*, 40% of the variation in allelic richness could be explained by latitude (Butcher *et al.* 2009), supporting the suggestion that the southern populations originated from the northern red gums (Hope 1994). The non-linear / non-radial migration patterns and evolutionary history of *E. globulus* (Freeman *et al.* 2001, see also 1.4.2 below) and the geographical barriers to dispersal that divide its distribution (the Murray Darling Depression, the highlands of the Great Dividing Range and Bass Strait, see Figure 1.1), could account for the fact that the relationship between distance from SE Forests and heterozygosity / allelic richness was not strong. Two lines of evidence give further strength to the hypothesis that the SE Forests region is the origin of the species complex though: a) the Neighbour-Joining tree, in which the SE Forests populations cluster centrally, and b) the capsule morphology in this region, which is assumed to be ancestral (Jordan *et al.* 1993), as seven-fruited umbels occur in other members of the subseries Globulinae (Slee *et al.* 2006). A continental origin of the *E. globulus* complex, followed by migration into Tasmania, is supported by chloroplast DNA evidence, as there was a high diversity of chloroplast haplotypes of the ancestral "C" clade, on continental Australia, and a widespread basal haplotype within this clade does not occur on the island of Tasmania (Freeman *et al.* 2001).

1.4.2 Evolutionary processes and migration patterns of E. globulus

This study has identified that the evolutionary processes of drift, selection and hybridisation have all had a role in shaping patterns of variation in *E. globulus*. The significant association between genetic and geographic distance is evidence that drift and/or mutation are strong processes in *E. globulus*. While the association between genetic distance and morphological distance among regions was also significant, it was small (r^2 = 0.38) and there were a number of key correlation breakers. For example, the South Gippsland region, of core *globulus* morphology, had molecular affinities to geographically close, but morphologically distant regions, such as Strzelecki Ranges, Buchan, and the three *pseudoglobulus* regions. The presence of these correlation breakers suggests that selection for large capsules over a steep environmental gradient

has occurred in the Gippsland region (Jones *et al.* 2002). Conversely, close morphological affinities but large genetic distances between 1-Mt Bryan and the other *bicostata* regions (region 2 to 7) is evidence not only for genetic drift but also selection maintaining the *bicostata* capsule morphology despite migration over large geographic distances. The lack of association between morphological and geographic distance is further evidence for selection on capsule morphology traits. However, as these samples were taken from natural populations rather than common environment field trials, the effect of environment cannot be discounted. Capsule morphology, however, does seem to have a strong genetic basis, since it appears to be maintained in field trials of the four taxa and their intergrades (pers. obs.) and there is a genetic basis to variation in capsule size at the subrace level in *globulus* (McGowen 2007).

Continuous clinal variation in microsatellite markers from the SE Forests region, west through Victoria into the 2-Mt. Cole region, with no close genetic links between East Gippsland populations and those on the inland side of the Great Dividing Range, strongly suggests that *E. globulus* dispersed westwards via a route on the coastal side of the Great Dividing Range, with the Great Dividing Range providing a strong barrier to dispersal. Genetic links between Mt Cole and populations both on the inland and coastal side of the Great Dividing Range, but a lack of evidence for recent hybridisation at Mt Cole, indicate Mt Cole may have been an intermediate population as the species migrated to the inland side of the Great Dividing Range, where it differentiated into *bicostata.* The *bicostata* at Mt. Bryan was the purest *bicostata* region in the STRUCTURE analyses, probably due to genetic drift, and also because it is not influenced by gene flow from geographically proximal but highly differentiated populations on the southern side of the Great Dividing Range. Indeed, this population must have been separated from the rest of the complex for a long time, as the marine incursions of the late Eocene to the mid Miocene and the aridity of the Murray Darling Depression since the start of the Oligocene (35 000 ya) would have meant this region was unable to sustain *E. globulus* populations, and the possibility of long distance seed/pollen dispersal over 700 km is remote.

The newly differentiated *bicostata* may have migrated north along the inland side of the Great Dividing Range. Contact with *maidenii* possibly occurred in the northern part of

the *bicostata* distribution, as the *bicostata* population at the 6-Jenolan region has molecular tendencies towards *maidenii*. The presence of a fairly pure *bicostata* population north of Jenolan (at 7-Wollemi) and evidence from the Neighbour-Joining tree (in which *maidenii* populations have closest links to *pseudoglobulus*, which in turn have affinities to *bicostata*, rather than *maidenii* having affinities to *bicostata* via the Jenolan population) argue against the hypothesis that *maidenii* differentiated into *bicostata* via the northern route (i.e. via Jenolan). Secondary contact between *maidenii* and *bicostata* (which had differentiated from *pseudoglobulus*) is a more parsimonious solution, and *E. globulus* therefore could be an example of a ring species (Endler 1977). Whether or not Jenolan is a recent or ancient hybrid zone could not be established in this study, as insufficient sampling meant that an association between morphological and molecular affinities and reliable measures of heterozygosity could not be determined. The 16-Omeo region, which was originally collected as *bicostata-pseudoglobulus* intergrade in the field and classified as *bicostata* intergrade by Jordan *et al.* (1993) is essentially part of the *bicostata* gene pool on the inland side of the Great Dividing Range, as evidenced by its morphometric and microsatellite affinities to *bicostata*.

The close microsatellite affinities among western Tasmanian populations, 21-King Island and 27-Phillip Island regions (and, to a lesser extent, Otways populations) and separate close affinities between eastern Tasmanian populations and 33-Furneaux and Wilson's Promontory populations, indicate that the Bass Strait has provided an incomplete barrier to gene flow, consistent with the presence of land bridges that linked north-eastern Tasmania and South Gippsland via the Furneaux Group (eastern route), and north-western Tasmania and the Melbourne region through King Island (western route) (Jackson 1999). Evidence from cpDNA studies also showed a western link between the Otways, King Island and Western Tasmania, suggesting this was the migratory route taken by the species into Tasmania, where it may have differentiated into the large-fruited *globulus*. However, there are no shared chloroplast haplotypes between north-eastern Tasmania and the Furneaux Group, arguing against seed mediated migration between mainland Tasmania and Victoria via an eastern route (Freeman *et al.* 2001). This pattern has also been observed in other eucalypt species (McKinnon *et al.* 1999, Marthick 2005). Studies of the distribution of nuclear markers such as microsatellites (this study and Steane *et al.* 2006) and the cinnamoyl coA reductase gene

(McKinnon *et al.* 2005) in *E. globulus* have provided evidence of nuclear links between populations in the Furneaux Group and Victoria which can be reconciled by pollen flow, but not seed flow, through this eastern route. This is supported by the presence of large fruited *globulus* on the coastal plains of southern Victoria which could have originated from ancient pollen swamping from Tasmanian *globulus* through the Furneaux Group into Victoria (i.e. secondary intergradation). Convergent selection for large capsules could have also occurred in these dry coastal areas as, for example, a mechanism to protect seeds during fire (see 1.4.3.4).

The presence of, for example, *maidenii*-type microsatellites underlying *pseudoglobulus* morphology in the Nadgee region, and *maidenii*-type microsatellites underlying *bicostata* morphology at Jenolan, as well as many disjunct populations across the range of the species complex, indicates that the taxa have undergone range expansion and contraction through time with changes in climate, glacial activity and sea levels associated with glacial cycles.

1.4.3 Origin of intergrade populations of the E. globulus *complex*

While the *E. globulus* gene pool is spatially structured, as evidenced by the AMOVA and the moderate and significant F_{ST}, the STRUCTURE and morphometric analyses showed that there is a continuum between these groups in the *E. globulus* gene pool. Many of the intergrade populations (discussed below on a case-by-case basis) are intermediate, with a genetic contribution from the core taxa, and have high morphological variability, which could indicate hybrid origin. Variability in capsule morphology in the intergrade populations led Jordan *et al.* (1993) to favour recent secondary intergradation as the origin of the intergrade zones, while Kirkpatrick (1975) suggested that the intergrade populations arose mostly through primary differentiation. Evidence from neutral microsatellite loci can shed further light on this, as a zone of recent secondary intergradation would be expected to have increased microsatellite diversity due to gene flow from differentiated populations (Rieseberg and Wendel 1993) and correlated morphological and molecular affinities, as was the case in the *Quercus crassifolia* - *Q. crassipes* hybrid zone in Mexico (Tovar-Sanchez and Oyama 2004).

1.4.3.1 11-Alfred-Nadgee intergrade region

The putative hybrid zone along the Cann Valley Highway had *maidenii*-like individuals in the north, *pseudoglobulus*-type individuals in the south and morphological intermediates between the two. However, the lack of correlation between morphological and molecular traits in the intermediate individuals, and the slightly lower diversity parameters in this locality compared with the putative parental localities, argue against recent hybrid origin. Instead, these intermediate individuals probably arose by primary differentiation as *E. globulus* moved from the SE Forests region into Gippsland and differentiated from *maidenii* inland to *pseudoglobulus* on the coast. The locality at the northern end of the Cann Valley Highway (locality 28) is the southern extent of the *maidenii* gene pool, along with 29-Wroxham and 31-Maramingo Creek localities. The 32-Alfred NP and 30-Nadgee localities have morphological affinities to *pseudoglobulus* with *maidenii* genes underneath, but there is no evidence of *maidenii* morphology with *pseudoglobulus* genetic affinities, further evidence that *maidenii* is the ancestral taxon of the complex.

1.4.3.2 15-Buchan intergrade region

It is likely that the intermediacy of the Buchan localities arose through primary differentiation as *E. globulus* migrated westward, rather than by recent hybridisation, as there was no significant association between morphological and molecular affinities and the Buchan intergrade region did not have higher diversity than the core *pseudoglobulus* populations in East Gippsland, which would be one of the putative parental populations based on morphology. There was geographic clustering of *pseudoglobulus* and *bicostata* morphotypes, the most obvious example being at the 45-Cutts Creek locality where all individuals had strong *bicostata* morphological affinities. This could be due to selection, but there is no obvious ecological difference between 45-Cutts Creek and the rest of the Buchan localities, which have morphological affinities to *pseudoglobulus*, *bicostata* and, in the case of one individual, *maidenii*.

1.4.3.3 17-Mitchell River intergrade region

Like the 15-Buchan region, the 17-Mitchell River region is also likely to have arisen through primary differentiation as *E. globulus* migrated westward. Individuals had microsatellite and morphological affinities to *globulus*, *bicostata* and *pseudoglobulus* but with no correlation between morphological and molecular traits. There was a correlation between *bicostata* morphology and *bicostata* molecular affinity but individuals had little morphological affinity to this taxon. There was a weak signal that *pseudoglobulus*-like morphology and microsatellites were correlated in a few individuals but this is unlikely to account for the overall intermediacy of the region. Mitchell River did not have higher levels of diversity than pure populations of *pseudoglobulus* and *bicostata*, also arguing against recent hybrid origin.

1.4.3.4 18-Strzelecki Ranges intergrade region

In the 18-Strzelecki Ranges region, the correlation between *globulus*-type morphology and genotype, but the lack of correlation between morphological and molecular affinities to other taxa, could be evidence for recent secondary contact between differentiated *globulus* in the 24-South Gippsland region and a zone of primary intergradation in the Strzelecki Ranges. The individuals with strong molecular and strong morphological affinity to *globulus* were randomly distributed rather than geographically clustered, and most individuals in this region had molecular affinities to *globulus*, indicating that the Strzelecki Ranges region is part of the Victorian *globulus* gene pool while selection for large capsules has occurred on the drier coastal plains of South Gippsland. Large capsules are thought to have a selective advantage in dry, fire-prone areas to insulate and protect seeds during fire (House 1997) and may also offer protection from bird predation (Clifford and Drake 1981). In *Hakea*, seed from large fruited species tend to be more viable than species with thin fruit walls (Bradstock *et al.* 1994). There is also a trend for the core *globulus* subraces (i.e. with larger capsules) to have heavier seeds (McGowen 2007) and this may convey a selective advantage in harsher environments. The reason that *globulus*-type capsules have also arisen by convergence in parts of the Strzelecki Ranges, but only in individuals with more *globulus*-like microsatellite genotypes, is difficult to explain on the basis of local selection. Alternatively, these *globulus*-like individuals in the Strzelecki Ranges could have been the original germplasm that

migrated to the coast of South Gippsland, with the large capsule morphology conferring a selective advantage in the region currently occupied by pure Victorian *globulus*.

1.4.3.5 The role of hybridisation in *E. globulus* intergrade populations

This combined analysis of morphological variables and neutral markers has shown that the *E. globulus* intergrade populations are unlikely to be of recent hybrid origin. Instead, they are likely to be the result of either primary differentiation or ancient hybridisation. The two latter evolutionary scenarios are difficult to distinguish using nuclear markers, but as cpDNA is non-recombinant as well as maternally inherited in most angiosperms (Birky 1995), it can sometimes be used to determine the occurrence of and direction of hybridisation. A hybrid population would contain two distinct parental chloroplast genomes rather than a single intermediate haplotype in a zone of primary differentiation. However, while the chloroplast is maternally inherited in *E. globulus* (McKinnon *et al.* 2001b), numerous studies of cpDNA variation in south-eastern Australia have shown that cpDNA variation among species is correlated with geography rather than species boundaries (Jackson *et al.* 1999; McKinnon *et al.* 1999; McKinnon *et al.* 2001a; McKinnon *et al.* 2004b; Steane *et al.* 1998) presumably as a result of hybridisation, introgression and chloroplast capture. The potentially extensive transfer of chloroplast types among species both within and beyond the *E. globulus* complex makes it difficult to use chloroplast DNA to determine the direction of gene flow in putative ancient hybrid zones at lower taxonomic levels. Indeed, a preliminary study of cpDNA variation in the Buchan intergrade region compared with pure *bicostata, pseudoglobulus* and *maidenii* was inconclusive due to high levels of haplotype diversity in the species complex (J. Bloomfield, UTAS unpubl. data).

The lack of clear genetic evidence for recent hybridisation as an evolutionary force shaping the patterns of diversity in the *E. globulus* complex is similar to the case of the *E. populnea-E. brownii* species complex where a genetic study indicated that this species complex was not a result of hybridisation (Holman *et al.* 2003). The authors went on to claim that a lack of evidence for hybridisation in the single case of *E. populnea-E. brownii* was evidence against the evolutionary significance of hybridisation in the genus *Eucalyptus* in general. However, accumulating evidence from cpDNA studies suggests that introgressive hybridisation has occurred between many

eucalypt species and has been a major evolutionary process at least in Tasmania, though perhaps not in Western Australia (Byrne 2008).

While the evidence suggests that much of the intermediate zone is differentiated by selection and isolation by distance, there is some evidence that secondary contact has occurred in various parts of the range. *Eucalyptus globulus* could be a ring species.

1.4.4 Conservation status of small and/or isolated populations

In some geographically or ecologically isolated populations, genetic drift and inbreeding have resulted in reduced genetic diversity and marked molecular divergence from the rest of the gene pool. This has implications for conservation planning.

1.4.4.1 Mt Bryan, South Australia

The population of *bicostata* found on private property at 1-Mt. Bryan in South Australia is around 700 km from the nearest other native population of the species (located at 2-Mt Cole in western Victoria) and is the only population west of the Murray-Darling drainage system. While *bicostata* is widespread through NSW and Victoria, it is listed as vulnerable in South Australia, since the Mt Bryan population is the only one remaining in this state (South Australia National Parks and Wildlife Act 1972). The population is of particular interest not only because of its isolation, but also because many of the trees at the site regenerate by lignotubers, resulting in clonal individuals up to ten metres in diameter. Previous work based on the rate of lignotuber growth suggested that one of the trees in this population was thousands of years old, confirming that it is a natural stand and not one planted by pastoralists in the last 200 years (Vaillancourt *et al.* 2001). On the basis of morphological and molecular markers, the Mt Bryan population is clearly part of the *bicostata* gene pool, but it is genetically distinct from other populations, and is therefore of high conservation value. The genetic divergence of this population supports the hypothesis that it has been geographically isolated for probably thousands of years, and genetic drift has occurred, resulting in differentiation from the rest of the *E. globulus* gene pool. The low levels of diversity at this site can be attributed to genetic drift resulting in allelic fixation, rather than the effects of inbreeding, as the *F* value, while positive, is no higher than the mean across

other populations of *bicostata*. Many of the trees that were sampled in this study could be thousands of years old, so that the genetic effects of inbreeding, usually associated with small disjunct populations, were not detected. The trees at this location were recently the focus of national attention due to their proximity to a proposed wind farm site, and information from this study was provided to ensure that the significance of the trees was considered in the wind farm planning process.

1.4.4.2 Nullo Mountain (Wollemi) region, New South Wales

The northernmost population of the *E. globulus* gene pool at Nullo Mountain State Forest (7-Wollemi) was clearly part of *bicostata* at the morphological and molecular level. However, like the 1-Mt Bryan population, it was a genetic outlier and therefore also of high conservation value. Also in similarity to the stand at Mt Bryan, it had low levels of diversity but a similar fixation index (F) to other *bicostata* populations, indicative of drift rather than inbreeding in this isolated population.

1.4.4.3 Jenolan Caves, New South Wales

The 6-Jenolan population of *bicostata* had molecular affinities to *maidenii*, which could be due to secondary intergradation. Unfortunately the sample size for morphological analysis was too small to determine the correlation between morphological and molecular affinities and identify the potential evolutionary processes involved in shaping such patterns of variation. Despite the small size of this population, and its isolation from the rest of the *bicostata* gene pool, it has moderate levels of genetic diversity and as it is currently well reserved as part of Jenolan Caves Reserve.

1.4.4.4 Mt Cole, Victoria

The disjunct population of *bicostata* at 2-Mt Cole in western Victoria is a key population that links the *E. globulus* gene pools on either side of the Great Dividing Range (see 1.4.2). While most individuals had morphological affinities to *bicostata*, it cannot be considered true core *bicostata* as there were only two trees sampled that had affinities to *bicostata* in both morphology and microsatellite genotype. While this population is disjunct, its large size has probably buffered it against the effects of drift as it has similar levels of diversity to other *bicostata* regions.

1.4.4.5 Lerderderg Gorge, Victoria

The small population at 14-Lerderderg Gorge in Victoria, once described as a separate taxon (*E. stjohnii*; the history of this taxon is described in Kirkpatrick 1974), is currently classified as *pseudoglobulus* even though it is 300 km west of other populations of *pseudoglobulus* in East Gippsland. The morphometric analysis showed that it deviates slightly from the core *pseudoglobulus* of the type localities in East Gippsland. Furthermore, it does not have genetic (microsatellite) affinities to the other *pseudoglobulus* populations, but is admixed, similar to nearby populations at Mt Cole and the Otways (regions 2, 19 and 20). The *pseudoglobulus*-like morphology of this region is possibly the result of convergence in capsule morphology towards the *pseudoglobulus* type. This highlights the problem of using highly variable morphological traits as taxonomic traits.

Despite its small size and isolation from the rest of the *pseudoglobulus* gene pool, the Lerderderg population has moderate levels of genetic diversity. It is probably a remnant from when *E. globulus* was more widespread through what is now Melbourne, prior to clearing. This would explain the genetic affinities of the Lerderderg population to populations at Mt. Cole and the Otways, all of which have microsatellite affinities to each other despite having divergent morphology. As the trees sampled in this study were all old, it may be too soon to detect the genetic effects of small population size and isolation. This interesting population is currently adequately reserved in Lerderderg State Park.

1.4.4.6 Nadgee River, New South Wales

The locality sampled at Nadgee River in NSW (locality 30), originally collected as *pseudoglobulus-maidenii* intergrade, had lower genetic diversity than other localities in the highly diverse SE Forests region and formed a unique cluster in the Victorian STRUCTURE analysis at $K = 6$, but its morphological affinities lay with *pseudoglobulus*. This small intergrade population is coastal, while the closest *maidenii* and p-m intergrades occur inland, apart from a phantom hybrid population at Mallacoota around 25 km away (with *E. cypellocarpa*, Kirkpatrick *et al.* 1973, not sampled in this study). The population is currently reserved in Nadgee Nature Reserve which is

sufficient to ensure this morphologically, genetically and geographically outlying population is conserved.

1.4.4.7 Isolated populations of *globulus* and its intergrades

There was evidence of molecular divergence and/or reduced genetic diversity in the small, disjunct populations of *globulus*, such as those at Wilson's Promontory and King Island. The genetically distinct Wilson's Promontory regions (25-Tidal River and 26-Wilson's Promontory) are currently reserved in a National Park, and are not under threat. The King Island populations (region 21) are mostly unreserved and part of remnant stands along watercourses on private land, but they are not genetically distinct; instead, they are part of the gene pool of small disjunct populations on the west coast of Tasmania, with genetic links to Victorian populations in the Otways and Phillip Island.

1.4.5 Taxonomic status of the E. globulus *species complex*

Three of the subspecies (*globulus*, *bicostata* and *maidenii*) are clear taxonomic entities: they are differentiated on the basis of capsule morphology and neutral molecular markers, and their cores are fairly uniform at both the molecular and the morphological level. However, the limited distribution of *pseudoglobulus* compared to the other core taxa, the high morphological variability within the core distribution of *pseudoglobulus*, the fact that it is the least differentiated taxon in terms of pairwise F_{ST} and Nei's genetic distance, and its failure to form a distinct group in the microsatellite analysis until $K = 7$, and, to some extent, in the morphological analysis, implies that it is not a meaningful taxonomic unit. Instead it appears to be part of a widespread intergrade population which extends across Victoria on the southern side of the Great Dividing Range.

As there is continuous morphological and molecular variation between the taxa of the *E. globulus* complex, and the intergrade populations occupy a large proportion of the gene pool, *bicostata*, *maidenii* and *globulus* are best treated as subspecies of *E. globulus* rather than separate species of a species complex.

There is clear evidence for selection on capsule morphology characters, which are the key taxonomic traits in *E. globulus*. The existence of regions with close molecular

affinities but highly divergent morphology (e.g. 18-Strzelecki Ranges *versus* 24-South Gippsland) as well as regions with similar morphology but differences at the molecular level (e.g. 14-Lerderderg *versus* 13-Lakes Entrance) suggests that capsule morphology is controlled by just a few genes, and while capsule traits are useful for differentiating the core taxa, they are not reliable for determining the genetic affinities of intergrade populations. Rather, geography appears to be the best predictor of genetic affinities in the *E. globulus* complex.

1.4.6 Practical implications for E. globulus *gene pool management*

This study, combining both morphology and neutral genetic markers, provides key baseline data for seed transfer guidelines in the *E. globulus* species complex. There have been cases where provenances such as Jeeralang (locality 55) have been labeled as *bicostata* due to their morphological similarity, but this analysis reveals that these populations are actually highly divergent. Guidelines based on evolutionarily consistent groupings, using a combination of neutral DNA and morphological evidence, will be more accurate. The STRUCTURE groups mapped in this study will provide a useful tool for choosing the appropriate germplasm for pulpwood plantations, breeding programs and conservation plantings and guiding seed transfer.

2 VARIATION IN FLOWER OPENING TIME AMONG RACES OF *E. GLOBULUS* SSP. *GLOBULUS*

2.1 Introduction

2.1.1 Flowering time as a barrier to gene flow in plants

Flowering time asynchrony is a major barrier to gene flow in plants (Levin 1978). While geographic barriers provide significant barriers to gene flow in *E. globulus* (see Chapter 1), annual flower opening time in *E. globulus* is also an important determinant of gene flow between natural populations, and also influences the tendency of the species to hybridise with co-occurring species. Eucalypts are known for their propensity to hybridise but flower opening time can provide a reproductive barrier between sympatric otherwise interfertile species (Griffin *et al.* 1988).

Annual flower opening time is one of the major factors that influences the potential for gene flow between plantations and native forest (Potts *et al.* 2001). *Eucalyptus globulus* ssp. *globulus* (hereafter referred to as *globulus*) is a major plantation species in southern parts of Australia and is planted both within and outside its natural range (Barbour *et al.* 2008b). An understanding of the genetic and environmental control of flower opening time is important to minimise the potential for gene flow between plantations and native forests.

As *globulus* is the most widely planted hardwood species in temperate parts of the world, there is significant demand for seed for plantations. Seed orchards are being used increasingly to produce genetically improved *globulus* seed, and the use of seed orchard seed rather than that collected from natural populations can provide significant genetic gains (Eldridge *et al.* 1993). An understanding of the timing of floral initiation and flower opening among elite genotypes in seed orchards is required for successful management of such orchards. For example, the timing of application of the GA inhibitor paclobutrazol requires knowledge of the flower bud initiation time because the hormone needs to be present in the apical tissue in time to effect the seasonal switch from vegetative to reproductive growth (Moncur *et al.* 1994). Annual flower opening

time is also a trait of interest because an overlap in flower opening time among trees is desirable when producing open pollinated (OP) seed, which is far cheaper than producting seed by hand-pollination. An ability to predict flower opening time based on the genotypes deployed and climate, would assist in site selection for optimal pollination among genotypes.

2.1.2 Genetic variation in flower initiation and opening time in globulus

Across the range of *globulus* in south-eastern Australia there is natural variation in flower opening time (Figure 2.1), and common environment field trials have shown that this trait has a genetic basis with a difference in peak flower opening time of up to six months among genotypes and races of *globulus* (Gore and Potts 1995, Apiolaza *et al.* 2001). There is, however, noticeable year-to-year variation in the onset and duration of flower opening (P.L. Gore pers. comm.), which indicates that climate may also be involved and that flower opening is not under photoperiodic control. Indeed, there are indications that heat sum may play a role in the onset of flower opening in eucalypts (see 2.1.4). Surprisingly, however, trees from different races do not appear to vary in the timing of macroscopic appearance of flower buds, which occurs in spring in all provenances (P.L. Gore pers. comm.). Early flower opening genotypes therefore appear to have a shorter flower bud development time than later flower opening genotypes, which could be caused by slower flower bud growth rates in late flower opening genotypes, or a period of dormancy during flower bud development.

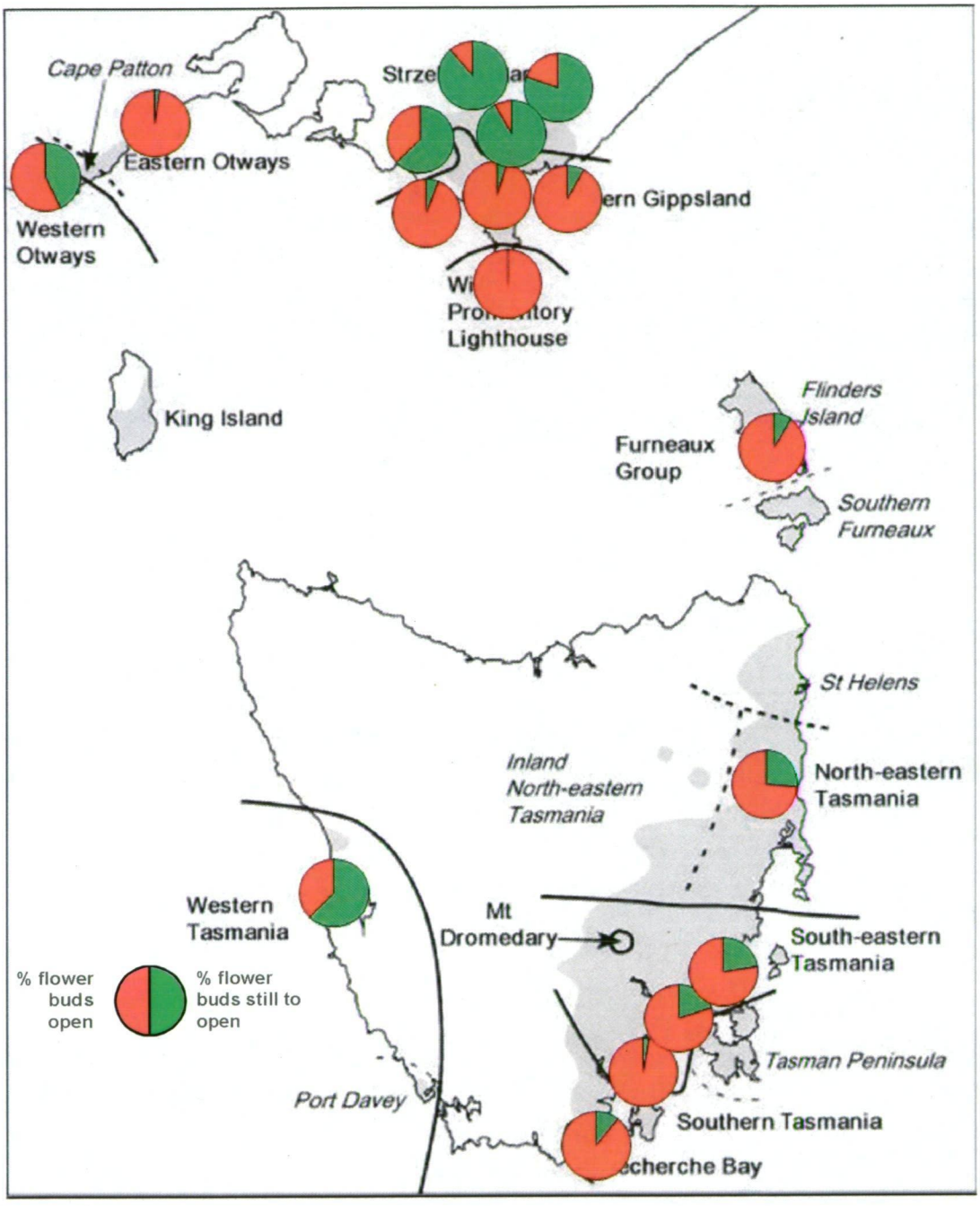


Figure 2.1. Flower opening surveyed in native stands of *globulus* and its intergrades in December 2006 (Potts and Vaillancourt, unpubl. data.). Races (Dutkowski and Potts 1999) are shown in bold and separated by solid lines, while sub-races are shown in italics, separated by dotted lines. Each pie chart is one population (figure provided by P. Tilyard). Pie charts with more red represent populations with more "early" genotypes while those with more green represent populations with more "late" flowering genotypes.

2.1.3 *Flower bud initiation and development in* globulus

Eucalyptus globulus ssp. *globulus* is heteroblastic, with a long juvenile, non-flowering phase. The juvenile form is distinct from the adult form: juvenile leaves are sessile, opposite and glaucous, and are carried on quadrangular stems, while adult leaves are petiolate, alternate and shiny green, on cylindrical stems. The species is a good model system for studying vegetative and reproductive phase change due to the marked contrast between the juvenile and adult forms. There are high levels of genetic variation in the timing of transition to adult foliage (Dutkowski and Potts 1999) and timing of first flowering (precocity, Chambers *et al.* 1997). These traits are independent and under moderate to strong genetic control, with individual narrow sense heritability (h^2) values of 0.2-0.9 and 0.4-0.6, respectively (Jordan *et al.* 1999). Nevertheless, flowers normally only occur in trees that are in their adult leaf phase.

The juvenile phase of *globulus* is followed by an annual cycling between vegetative and reproductive growth (Figure 2.2). It is evergreen, but over the winter months vegetative growth is limited. Following the annual spring flush of vegetative growth, floral buds are observed in leaf axils, though the flower buds probably initiate at the microscopic level several weeks before they are visible to the naked eye (6-8 weeks in *E. nitens*; Moncur *et al.* 1994) and even earlier at the molecular level. The inflorescences are usually single-flowered but occasionally have three flowers per umbel (see 1.1.2). During the early stages of flower development, the sepal primordia fuse to each other to form an outer operculum and the petal primordia fuse to form an inner operculum (Figure 2.3). In some eucalypt species these two operculae are fused to each other, but in *globulus* they are separate structures, as in other species from the subgenus *Symphyomyrtus* (Ladiges 1997). The entire umbel is enclosed by a pair of bracts which are shed early in flower development as the bud elongates, followed by the shedding of the outer sepaline operculum. The shedding of the inner (petaline) operculum to expose the stamens occurs once all parts of the flower bud are fully developed, from late autumn to early summer in *globulus* and is the process commonly referred to as "flowering". Here, the shedding of the inner operculum will be termed "flower opening" to avoid confusion with the flower bud initiation (FI) stage, as the latter corresponds to the stage normally referred to as "flowering" in the model annual plant *Arabidopsis*.

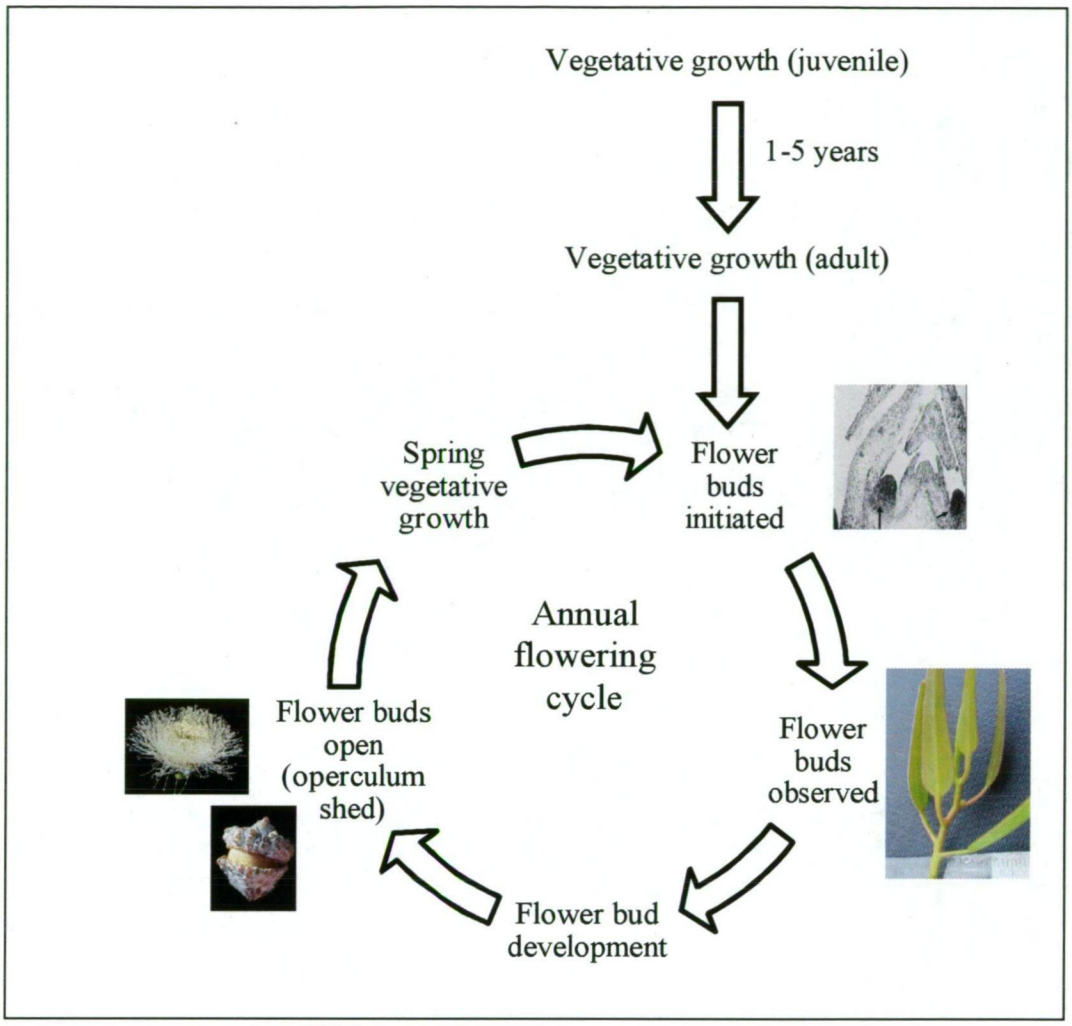


Figure 2.2. The transition to flowering in *globulus*. Microscope image is of *Metrosideros excelsa* (Sreekantan *et al.* 2004), and the arrows indicate flower bud primordia.

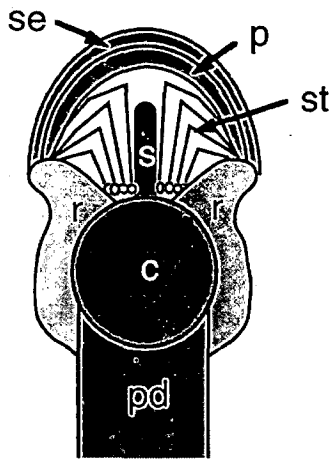


Figure 2.3. Schematic cross section of a typical eucalypt flower bud (from Southerton *et al.* 1998a). c, carpel; p, petal; pd, peduncle; r, receptacle; s, style; se, sepal; st, stamen. The outer and inner operculae are fused sepals and petals, respectively.

2.1.4 Plant responses to heat sum

Temperature is an important factor that influences the developmental rate of many organisms, especially in plants, where a certain amount of heat is required for a particular developmental process to occur. The measure of accumulated heat above a base temperature is known as heat sum, or thermal time, and is measured in degree days or degree hours. Heat sum has been implicated in the timing of many developmental processes, especially in agricultural crops, and trees for fruit production and forestry (reviewed in Trudgill *et al.* 2005). In tree systems, heat sum models have been used to analyse the effects of temperature on developmental processes such as spring bud burst in temperate and boreal trees, and seasonal flower opening. For example, most of the year-to-year variation in peak flowering of *Pinus elliottii*, *P. palustris*, *P. taeda* and *P. echinata* can be attributed to heat sum (Boyer 1978). Using 19 years of flowering observations for *P. palustris* and six years of data for the other species, the average deviation from observed to predicted peak flowering was four days or less (Boyer 1978). Heat sum models have also been used to predict flower opening time in other trees such as almond (Rattigan and Hill 1986, 1988), Norway Spruce (Nikkanen 2001), *Quercus* (Garcia-Mozo *et al.* 2006; Rodriguez-Rajo *et al.* 2005), *Alnus glutinosa* (Gonzalez-Parrado *et al.* 2006), *Prunus padus* and *Tilia cordata* (Thompson and Clark 2006), olive (Perez-Lopez *et al.* 2008) and nine shrubs and trees in the Mediterranean (Spano *et al.*

1999). In some cases heat sum alone accounted for the year-to-year or site-to-site variation in flower opening, but in other cases a model combining chilling units and heat sum was used.

There are indications that heat sum may play a role in the year-to-year or site-to-site variation in annual flower opening time in eucalypts. For example, the flower opening time of plantings of the exotic *Eucalyptus nitens* in Tasmania was earliest in years that had the greatest annual heat sum (Barbour *et al.* 2006). Flower opening was also later in high altitude than in low altitude plantations that were based on random genetic material, and the flower opening period was also shorter at high altitudes (Moncur *et al.* 1994, Barbour *et al.* 2006). Earlier flower opening at low altitudes has also been observed in other eucalypts (Pryor 1976). For example, in natural stands of *E. regnans*, flower opening commenced approximately two weeks earlier in lower altitudes than in high altitudes (Ashton 1975, Griffin 1980). Temperature has also been implicated in the site-to-site variation observed in melliferous (honey-producing) eucalypts, as northern (warmer) sites often flower earlier than southern (cooler) sites (Birtchnell and Gibson 2006).

Concerns about global warming have sparked interest in understanding the interactions between temperature and growth/phenology. Heat sum dependent processes are commonly incorporated into models to predict the potential effects of climate change. For example, Sykes and Prentice (1996) used growing degree days as part of a model to predict the range of tree species in the forests of northern Europe under different climate change scenarios. Other studies have already shown that changes in phenology are occurring in response to shifting temperature regimes: a meta-analysis of the phenology of 542 plant and 19 animal species across Europe over the last thirty years demonstrated a clear relationship between temperature and phenology, with evidence of advancing leaf unfolding, flowering and fruiting, and delays in leaf colouring and leaf fall in wild plants all across Europe (Menzel *et al.* 2006). On average, spring was advancing by 2.5 days per decade and leaf colouring and fall were delayed by 1.0 day per decade (Menzel *et al.* 2006). Clearly, there is a need to determine the relationship between temperature and phenology in eucalypts. Such information is critical in predicting the effects of global warming on flower opening time, which will in turn affect nectar availability for

birds and insects and the likelihood of cross-pollination within or between natural and planted stands. If the climate change response is different between species, then the synchrony or asynchrony in phenology between species could be disrupted. The endangered swift parrot (*Lathamus discolor*), feeds on the nectar of *Eucalyptus globulus* flowers and the yearly reproductive success of the parrot is thought to depend upon the flowering intensity of *globulus* (Hingston *et al.* 2004a). The relationship between *globulus* flowering and swift parrot migration could be disrupted if the increase in temperature affects the phenology of the two species differently.

2.1.5 Aims of this study

The objectives of this study were to (i) determine the genetic control of flower bud initiation and flower bud opening time in *globulus*, and (ii) examine the relationship between climate (heat sum) and flower bud opening time. Macroscopic appearance of flower buds was monitored over two seasons, and flower bud opening time was monitored over six seasons, in a seed orchard comprising around 700 trees representing four races and 89 genotypes of *globulus*. Information from this study will be valuable in the development of predictive models of flower opening time in a seed orchard, plantation or native stand, based on knowledge of the genotype or provenance deployed, and the climate of the local area.

2.2 Materials and Methods

2.2.1 Seed orchard

Trees used in this study were located in the seedEnergy Pty. Ltd. seed production orchard located at Cambridge, Tasmania at 42°48'26"S, 147°25'52"E and 40m elevation. All genotypes were selected based on breeding value for pulpwood production from the Southern Tree Breeding Association (STBA) breeding population of *globulus* and represented six of the races defined by Dutkowski and Potts (1999): Eastern Otways, Western Otways, Strzelecki Ranges, Furneaux, South-Eastern Tasmania and Southern Tasmania. As some of these races are not highly differentiated at the molecular level (Chapter 1), these races were pooled into four key lineages: East Coast South (South-Eastern Tasmania and Southern Tasmania), Otway Ranges (Eastern Otways and Western Otways), Strzelecki Ranges and Furneaux Group. Clones (i.e. ramets of each genotype) had been grafted before planting, and later treated with paclobutrazol, a GA inhibitor that reduces internode length and increases the intensity of, but does not affect the timing of, flower bud initiation in eucalypts (Hetherington *et al.* 1992; Hasan and Reid 1995). Genotypes were arranged in multi-ramet line plots with one to 52 plots per genotype. The plots were effectively randomly distributed throughout the orchard and varied in size from one to 22 ramets per plot. The distribution of genotype from different races within the orchard was also effectively random.

2.2.2 Flowering surveys

Surveys were undertaken at least every two weeks during the flowering season of *globulus*, which generally lasts from May to February at the Cambridge seed orchard. Flower opening time (FT) was measured by recording the percentage of unopened flower buds (p) on the tree at each survey, with p decreasing from 100% to 0% during the time of scoring. Flower bud initiation (FIobs) was recorded as the survey date (number of days after April 1 each year) at which flower buds were first macroscopically visible. FT was scored in the flowering seasons starting in 2003, 2004, 2005, 2006, 2007 and 2008, but less trees were surveyed in 2003, and FT was not

surveyed for the entire length of the 2003, 2005 and 2008 seasons (Table 2.1) as these surveys were conducted as part of the orchard's operational program. FI was scored in the 2006 and 2007 seasons. Date of flush of vegetative growth (Veg) was also recorded in 2007.

Table 2.1. Number (*n*) of families, genotypes and individuals of *globulus* surveyed for flower initiation (FIobs) and three phases of flower opening (FT1, FTpeak and endfl) from each race each year in the seedEnergy Cambridge orchard. *n* varies between years because not all individuals flowered each year. Races are FG, Furneaux Group; ECS, East Coast South; OR, Otway Ranges; SR, Strzelecki Ranges.

		FIobs			FT1			FTpeak			endfl		
Year	Race	*n* (families)	*n* (genotypes)	*n* (individuals)	*n* (families)	*n* (genotypes)	*n* (individuals)	*n* (families)	*n* (genotypes)	*n* (individuals)	*n* (families)	*n* (genotypes)	*n* (individuals)
2003	FG				8	10	17	8	10	17	8	10	17
2003	OR				2	4	5	2	4	5	2	4	5
2003	SR				5	5	7	5	5	7	5	5	7
2004	ECS				3	4	24	3	4	24	3	4	24
2004	FG				14	23	198	14	23	198	14	23	198
2004	OR				8	14	85	8	14	84	6	8	49
2004	SR				11	14	88	11	14	88	10	13	85
2005	ECS				3	4	11	2	3	5	1	1	1
2005	FG				11	18	115	9	15	81	6	9	20
2005	OR				3	3	18	3	3	15	1	1	2
2005	SR				4	4	40	4	4	19			
2006	ECS	3	6	23	2	4	20	3	5	21	3	5	21
2006	FG	13	21	169	9	16	105	11	19	124	11	19	124
2006	OR	9	16	102	8	15	91	8	15	97	8	15	97
2006	SR	11	13	81	10	13	76	10	13	81	10	13	80
2007	ECS	3	6	20	3	6	12	3	6	12	3	6	12
2007	FG	14	23	211	13	21	175	13	21	174	13	21	173
2007	OR	9	16	124	9	15	101	9	15	101	9	15	101
2007	SR	11	13	97	11	13	80	11	13	80	11	13	80
2008[1]	ECS				3	4	12	3	6	19	3	6	18
2008[1]	FG				13	21	168	13	21	199	13	21	198
2008[1]	OR				7	12	61	9	16	123	9	16	122
2008[1]	SR				8	8	53	11	13	93	11	13	92

[1] 2008 data includes both MSP and survey data combined.

The FT survey data were converted into three components for each tree, all measured in days from April 1 each year: day at which flowering commenced (FT1), day at which flowering terminated (endfl) and mean flowering time (FTpeak). FTpeak was the estimated day at which 50% of the tree was in flower. FTpeak was estimated by fitting a cumulative density function to the flowering data, expressed as 1-*p*/100 and calculating the 50th percentile (Apiolaza *et al.* 2001). A two-parameter Weibull function was fitted to the data based on PROC NLIN in SAS (SAS Institute; Version 9.1). This model was fitted separately for each flowering season. Flowering period (flperiod; i.e. endfl-FT1) was also calculated for each tree each year.

In 2008 the FT survey was incomplete, but records of mass supplementary pollination (MSP, Patterson *et al.* 2004) that had been undertaken on some trees were available. For each tree, the number of flowers pollinated per day was divided by the sum of buds pollinated over the whole flowering season to give the proportion of flower buds open per day. These proportions were converted into the percentage of unopened flower buds on the tree each day, to make it equivalent to FT survey data. A subset of trees (n=50) had FT1 scores using both survey and MSP data allowing a comparison between data sources. These measures were highly correlated (Pearson $r = 0.943$, $P < 0.0001$) but the means for survey (FT1) and MSP data (FT1msp) for this subset of trees were 189.7 and 196.02 respectively. Therefore, where FT survey data was missing, FTmsp data was substituted but with an adjustment to account for the differences in means (i.e. if FT1 missing then FT1 = FT1msp-6.32).

As flower bud initiation time could be measured as the time at which buds were either macroscopically or microscopically visible, and also because variation in this trait could be due to scoring inexperience in detecting new flower buds rather than true phenotypic variation (see 2.3.1), five different measures of flower initiation time were used, all measured from April 1 each year: (1) the day at which flower buds were first macroscopically visible (FIobs, only measured in 2006 and 2007); (2) the day at which flower buds were probably microscopically visible, calculated as FIobs-49 as in the closely related species *E. nitens* flower initiation occurs at the microscopic level 6-8 weeks before flower buds are macroscopically observed (Moncur *et al.* 1994) (FIobsMon, only calculated in 2006 and 2007); (3) estimated day of flower bud

initiation, based on the overlapping modes of FIobs in 2006 and 2007 (see Figure 2.5) (FIest, = day 220) and (4) FIest-49 (FIestMon, = 171); and (5) estimated day of flower bud initiation, based on the race means averaged across 2006 and 2007 (FIestrace, see 2.3.1: East Coast South = 230.85, Furneaux Group = 229.91, Otway Ranges = 227.69, Strzelecki Ranges = 219.94).

2.2.3 Flower bud development

To compare flower bud development in an early (4489, Furneaux Group) and late (5296, Strzelecki Ranges) flower opening genotype, bud length, bud width and shedding of the flower bract and inner and outer operculae were monitored on two ramets of genotype 5296 (ramets 12-34 and 21-02, "Late-3" and "Late-4" respectively) and one ramet of 4489 (ramet 13-28, "Early-2"). Ramet identifiers include the south-north coordinate followed by the east-west coordinate. As only one ramet of 4489 initiated flower buds in 2006, no further replicates could be used for this genotype. On each ramet, four branches with at least five flower buds each were tagged and measured every two weeks from 17th January 2007 until flower bud opening time. The branches faced due north, south, east and west.

2.2.4 Climatic data

Climatic data for the Hobart airport weather station (42°49'59.16"S 147°30'11.88"E, elevation 3m; 6.5 km from seed orchard) from October 2002 to February 2009 were provided by the Bureau of Meteorology, Hobart. Daily records obtained were: minimum and maximum temperatures (°C), mean temperature (°C, calculated by averaging the daily minimum and maximum temperatures), solar exposure (MJ/m^2, total amount of solar energy falling on a horizontal surface of unit area, derived from satellite imagery), and number of hours of bright sunshine. Temperature readings, recorded every minute, were also obtained and these were averaged per day to give a more accurate mean daily temperature than the average of the minimum and maximum daily temperature.

Heat sum is the linear accumulation of temperature above a base temperature (Trudgill *et al.* 2005). In this study two different base temperatures were used: four degrees, which is the approximate temperature at which growth and development becomes negligible in *E. nitens* (Moncur *et al.* 1994) and five degrees, identified as the base temperature in a study of *globulus* growth (Reed *et al.* 2003). For the period from FI (FIobs, FIobsMon, FIest, FIestMon or FIestrace) to FT (FT1 or FTpeak) in each ramet, each year, positive daily heat sum values above four and five degrees ("Heatsum4" and "Heatsum5" respectively), based on means derived from maximum and minimum daily temperatures, as well as positive daily heat sum values above four degrees based on minute data ("Heatsum4min") were summed. In addition, solar exposure ("Solar") and number of sunshine hours ("Sunshine") were summed for the period from FI to FT. Therefore, for each tree, there were five different climatic variables summed over ten different FI to FT time periods for the 2007 and 2008 seasons and six different FI to FT time periods in the 2003-2006 seasons.

2.2.5 Data analysis

2.2.5.1 Trait variation among races

Variation of each trait was analysed separately for each year as a mixed model using PROC MIXED in SAS (SAS Institute; Version 9.1), with race as a fixed effect, and genotype within race and plot within genotype as the random effects. Family effects were rarely significant as there was a low level of replication of genotypes within families so this level of hierarchy was excluded.

To obtain an overall estimate of the relative importance of each term in the model, the above model was fitted to the data treating all terms as random. In this case the estimated variance components were used to estimate the percentage of the total variance attributable to each term.

As the measurements on each individual tree across time were not independent, a repeated measures model was fitted with ramet within plot as the subject. The first year (2003) was excluded from this analysis as there were only three races represented in this

year. The full model could not be fitted for the FTpeak trait so the covariances in the error strata were dropped in this case.

2.2.5.2 Correlations among traits

Correlation analyses were used to test for associations between traits, using PROC CORR in SAS (SAS Institute; Version 9.1). Spearman correlations were calculated at the phenotypic (i.e. individual trees ignoring genotype in the analysis, $r_{phenotypic}$) and genotypic (genotype arithmetic means calculated with PROC MEANS, $r_{genotypic}$) levels for FIobs in 2006 and 2007, FIobs and Veg in 2007, Veg in 2007 and FT1 in 2008, FT1 and FTpeak within each year, FIobs 2006 and FTpeak 2007, FIobs 2007 and FTpeak 2008, and FTpeak among all years. Pearson correlations were calculated for the daily calculations/measurements of the five climatic variables.

2.2.5.3 Heritability

The broad-sense heritability (H^2) and its standard error were calculated for FIobs, FT1, FTpeak and flperiod, using the program ASReml (Gilmour *et al.* 1995). The model fitted year as a fixed effect and genotype, plot within genotype and their interactions with year as random effects. The model ignored race of origin and thus the genotype term included race and within race genetic effects. The broad-sense heritability (Falconer and MacKay 1996) was calculated by dividing the genotype variance component by the sum of all variance components for all random terms in the model, and included a pooled residual term.

2.2.5.4 Association between climate variables and flower bud development time

A subset of individual trees that had a complete five year series of FT1 or FTpeak records (n=79 and n=70 respectively) was selected to study the differences among years in the accumulation of climatic variables over ten different FI to FT time periods as outlined in 2.2.4. The differences between years was studied by fitting a one way repeated measures model with year as a fixed effect and estimating differences between least squared means following a Tukey-Kramer adjustment.

2.3 Results

2.3.1 Timing of flower bud initiation and vegetative flush

Flower bud initiation occurred nearly simultaneously in all clones, with no significant difference within each year in mean flower bud initiation time among races (Table 2.2, Figure 2.4). In 2007 the genotype effect was significant but not in 2006 (Table 2.2). Flower initiation in 2006 and 2007 were significantly correlated at the phenotypic level (i.e. individual trees ignoring genotype) and genotype level but the correlations were weak (Spearman $r_{phenotypic} = 0.26$, $P < 0.0001$, n=356; Spearman $r_{genotypic} = 0.22$, $P = 0.0062$, n=154). The phenotypic and genotypic correlations between FIobs and Veg in 2007 were positive and significantly greater than zero (Spearman $r_{phenotypic} = 0.62$, $P < 0.0001$, n=85; Spearman $r_{genotypic} = 0.52$, $P < 0.0001$, n=83) and the mean time of vegetative flush was around three days earlier than mean flower bud initiation time in this year (data not shown). The timing of vegetative flush (Veg) in 2007 was not correlated with the timing of first flowering (FT1) in 2008 (Spearman $r_{phenotypic} = 0.09$, ns, n=48; Spearman $r_{genotypic} = -0.05$, ns, n=65). Flower bud initiation time across the two years had a very low broad sense heritability (H^2) of 0.06±0.05 (Table 2.3).

Table 2.2. The fixed race F values (F), numerator and denominator degrees of freedom (df) for the test of race effect and significance (P) as well as random genotype within race Z value (Z) and significance for the flower bud initiation (FIobs), first flowering (FT1) and peak flowering (FTpeak) traits, calculated separately each year for the *globulus* seedEnergy Cambridge orchard.

	FIobs				FT1				FTpeak			
	Race		Genotype (Race)		Race		Genotype (Race)		Race		Genotype (Race)	
Year	F (df)	P	Z	P	F (df)	P	Z	P	F (df)	P	Z	P
2003					12.00 (2,16)	0.0007	1.89	0.0294	11.99 (2,16)	0.0007	0.07	0.4738
2004					15.00 (3,51)	<0.0001	4.27	<0.0001	17.43 (3,51)	<0.0001	4.58	<0.0001
2005					1.76 (3,25)	0.1804	2.45	0.0071	3.07 (3,21)	0.0503	2.22	0.0134
2006	2.45 (3,52)	0.0736	1.21	0.1125	2.79 (3,44)	0.0517	3.01	0.0013	6.33 (3,48)	0.0011	3.97	<0.0001
2007	2.56 (3,54)	0.0644	2.36	0.0092	12.02 (3,51)	<0.0001	3.97	<0.0001	10.38 (3,51)	<0.0001	4.40	<0.0001
2008					13.12 (3,41)	<0.0001	4.16	<0.0001	15.85 (3,52)	<0.0001	4.46	<0.0001

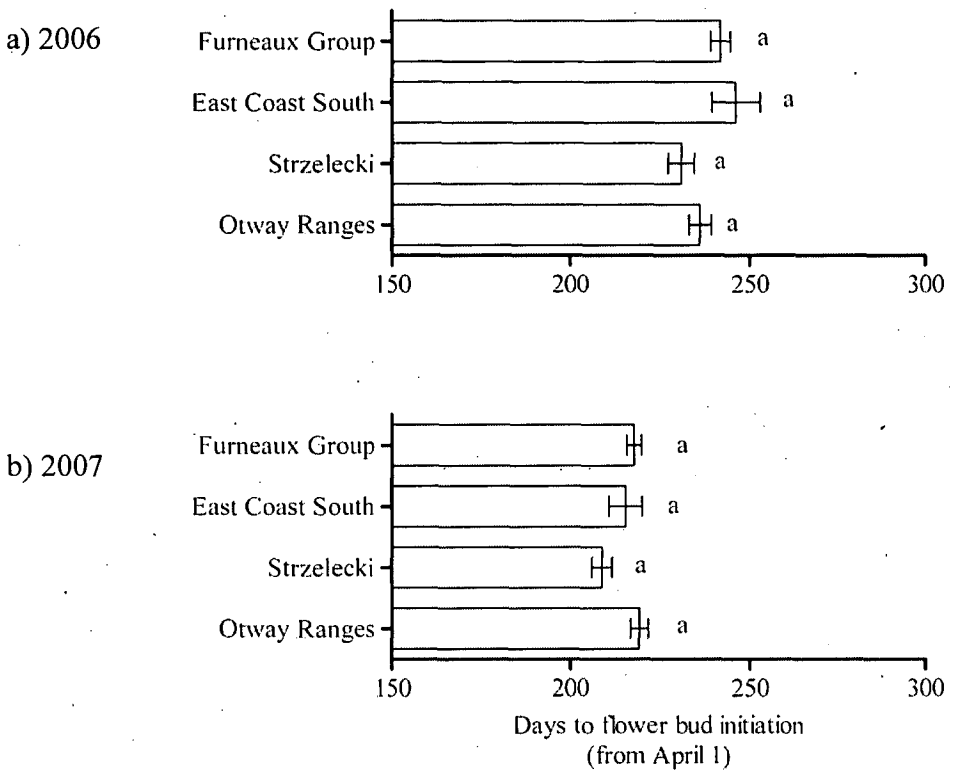


Figure 2.4. Least-square mean FIobs per race of *globulus* each year at the seedEnergy Cambridge orchard in a) 2006; and b) 2007. Results are from the analysis of variance presented in Table 2.2. Error bars represent one standard error, and race means sharing a letter within a year were not significantly different using the Tukey-Kramer test.

Table 2.3. Broad sense heritabilities (H^2) and their standard errors (SE) of flower bud initiation (FIobs), the two flower opening time traits (FT1, FTpeak) and flower opening period (flperiod) measured at the *globulus* seedEnergy Cambridge orchard.

Trait	H^2	SE
FIobs	0.06	0.05
FT1	0.69	0.05
FTpeak	0.77	0.04
flperiod	0.09	0.05

There was a highly significant difference in mean FIobs between the two years (Table 2.4) and a significant year x genotype interaction (Table 2.5). While the mean FIobs was later in 2006 (mean = 238.82) than in 2007 (mean = 215.38), this may have been an artefact of scoring inexperience in detecting new flower buds in the first year of FIobs surveys, as the mode day of FIobs was 220 in both years (Figure 2.5). In this combined analysis the race effect was slightly significant, as the Strzelecki Ranges race had slightly earlier flower bud initiation time (Figure 2.4).

Table 2.4. F-values (F), degrees of freedom (df) and significance (P) of fixed effects in the repeated measures model of flower bud initiation (FIobs), first flowering (FT1) and peak flowering (FTpeak) at the *globulus* seedEnergy Cambridge orchard.

	FIobs		FT1		FTpeak	
Effect	F (df)	P	F (df)	P	F (df)	P
Year	99.25 (1)	<0.0001	31.73 (4)	<0.0001	24.56 (4)	<0.0001
Race	3.56 (3)	0.0204	14.42 (3)	<0.0001	13.86 (3)	<0.0001
Year x race	1.37 (3)	0.2626	2.90 (12)	0.0012	5.35 (12)	<0.0001

Table 2.5. Z-values (Z) and significance (P) of random effects in the repeated measures model of flower bud initiation (FIobs), first flowering (FT1) and peak flowering (FTpeak) at the *globulus* seedEnergy Cambridge orchard.

	FIobs		FT1		FTpeak	
Effect	Z	P	Z	P	Z	P
Genotype within race	0.42	0.3364	4.39	<0.0001	4.69	<0.0001
Plot within genotype	2.40	0.0081	2.23	0.0128	3.70	0.0001
Year x genotype within race	1.95	0.0256	3.88	<0.0001	5.48	<0.0001

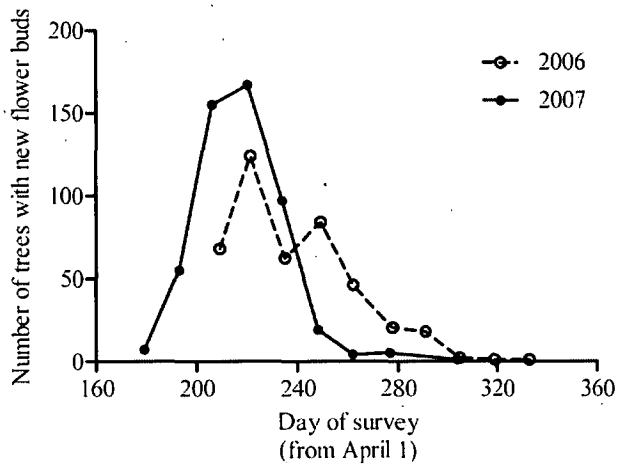


Figure 2.5. Number of trees on which flower buds were observed for the first time at each survey day in 2006 and 2007 at the *globulus* seedEnergy Cambridge orchard.

2.3.2 Timing of flower bud opening

The two measures of flower opening time (FT1 and FTpeak) were highly correlated each year (mean Spearman $r_{phenotypic} = 0.86$, mean Spearman $r_{genotypic} = 0.91$) and significantly greater than zero (Table 2.6). As survey dates were sometimes up to two weeks apart, the calculated FTpeak was potentially more accurate than FT1 and therefore FTpeak was mostly used in further analyses. FIobs (2006) and FTpeak (2007) were weakly positively correlated at the tree level (Spearman $r_{phenotypic} = 0.15$, $P = 0.0052$, n=344) but not at the genotype level (Spearman $r_{genotypic} = -0.04$, $P = 0.6258$, n=151). FIobs (2007) and FTpeak (2008) were positively correlated at the tree level (Spearman $r_{phenotypic} = 0.26$, $P < 0.0001$, n=427) and the genotype level (Spearman $r_{genotypic} = 0.23$, $P = 0.0036$, n=166).

Table 2.6. Phenotypic (ramet level) and genotypic (genotype level) Spearman correlation (r) between the FT1 and FTpeak traits each year at the *globulus* seedEnergy Cambridge orchard. Significance levels are indicated (*** $P < 0.001$) and number of observations is given in parentheses.

Year	Phenotypic r	Genotypic r
2003	0.84*** (29)	0.88*** (40)
2004	0.91*** (394)	0.97*** (125)
2005	0.76*** (120)	0.81*** (64)
2006	0.84*** (292)	0.92*** (128)
2007	0.93*** (367)	0.95*** (153)
2008	0.88*** (294)	0.91*** (131)

Each year there was a significant difference in mean flower opening time among races and sometimes among genotypes (Table 2.2, Figure 2.6). Family effects were rarely significant (data not shown) and so this level of hierarchy was excluded from subsequent analyses. Overall, all fixed and random effects were significant for the FT1 and FTpeak traits (Table 2.4, Table 2.5). Rank FTpeak was highly positively correlated across years both at the phenotypic level (mean Spearman $r_{phenotypic} = 0.72$, range 0.36-0.88, Table 2.7) and the genotypic level (mean Spearman $r_{genotypic} = 0.78$, range 0.31-0.93, Table 2.7). This could also be seen to some extent at the race level with the Furneaux Group and East Coast South races always flowering earliest and Otway Ranges and Strzelecki Ranges flowering latest each year (Figure 2.6, Figure 2.7). The tendency of the Strzelecki Ranges race to initiate buds slightly earlier than other races (see 2.3.1) cannot account for the variation in flower bud opening time, as the Strzelecki Ranges race flowers late, not early (Figure 2.6).

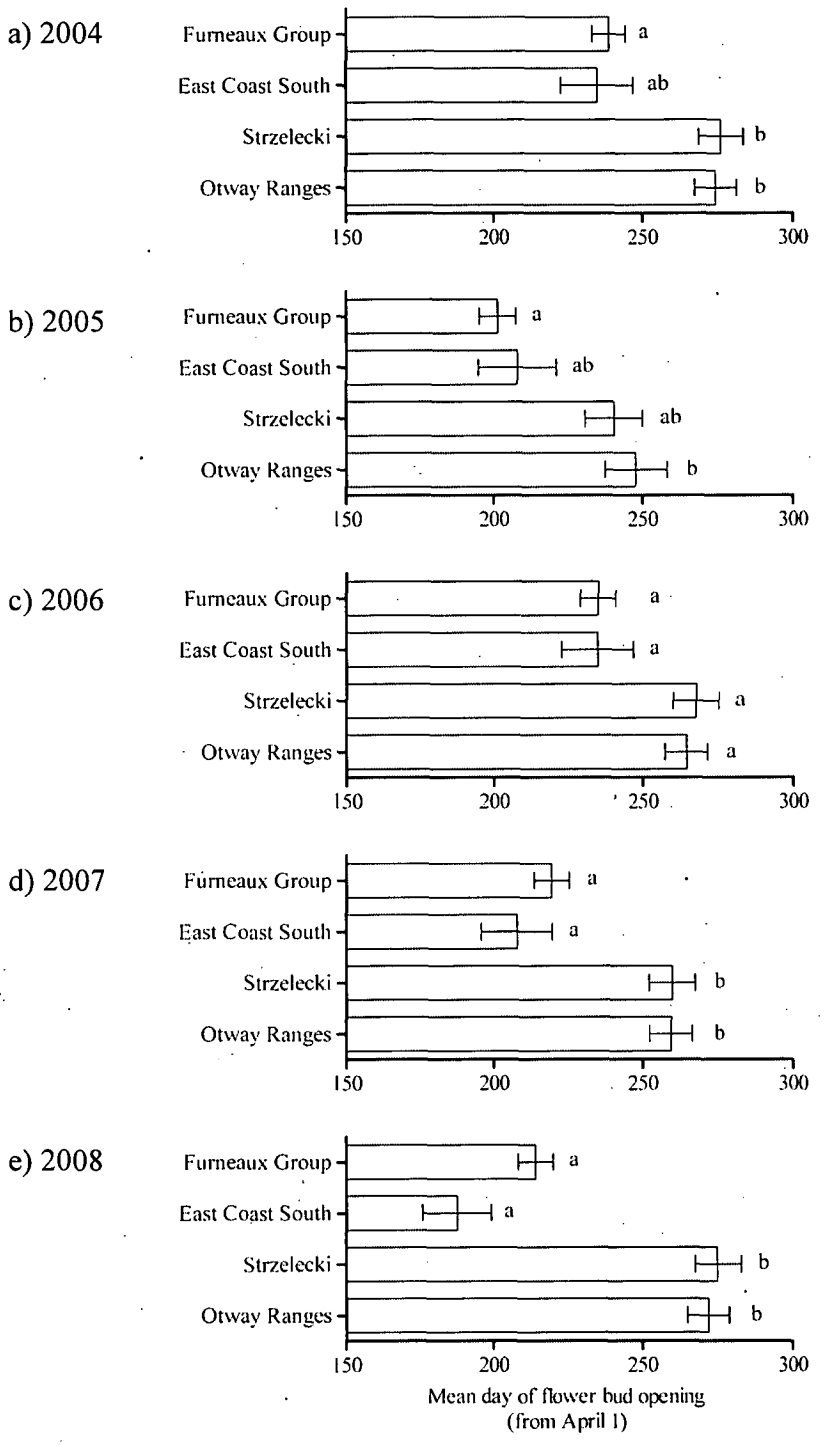


Figure 2.6. Least-square mean FTpeak per race of *globulus* each year at the seedEnergy Cambridge orchard in a) 2004; b) 2005; c) 2006; d) 2007; e) 2008. Results are from the analysis of variance presented in Table 2.2, which includes all flowering observations. Error bars represent one standard error, and race means sharing a letter within a year were not significantly different using the Tukey-Kramer test.

Table 2.7. Spearman correlations (r) of FTpeak across years at the ramet level (above diagonal) and the genotype level (below diagonal) at the *globulus* seedEnergy Cambridge orchard. Significance levels are indicated (*** P < 0.001; ns not significant) and *n* is given in parentheses.

	2003	2004	2005	2006	2007	2008
2003		0.66*** (26)	0.50ns (9)	0.70** (15)	0.82*** (22)	0.76*** (21)
2004	0.87*** (37)		0.69*** (122)	0.80*** (280)	0.86*** (314)	0.85*** (348)
2005	0.31ns (19)	0.79*** (55)		0.36*** (96)	0.68*** (127)	0.60*** (130)
2006	0.74*** (32)	0.87*** (106)	0.54*** (60)		0.85*** (286)	0.80*** (320)
2007	0.73*** (34)	0.92*** (116)	0.74*** (64)	0.90*** (125)		0.88*** (395)
2008	0.82*** (31)	0.93*** (113)	0.68*** (63)	0.89*** (128)	0.93*** (148)	

There was variation in the degree of flowering synchrony among races each year, as evidenced by the significant year x race interaction (Table 2.4). For example, in 2006 all of the races flowered late in the season (after day 200), and there was therefore a large degree of overlap in flowering period among races (Figure 2.7). However, in 2007 and 2008 the Furneaux Group and East Coast South races began flowering around 100 days earlier than the Strzelecki Ranges and Otway Ranges races, and there was less overlap between the early and late flower opening races (Figure 2.7). FT1 and endfl data in the 2003, 2004 and 2005 years were less reliable, because the surveys in these years were conducted as part of the orchard's operational program and therefore the surveys sometimes began after the start of the flowering season, or were abandoned before the end of the season, so data for these years were not graphed. In 2008 the assessment just missed the first few trees flowering, so minimum FT1 could not be plotted for Furneaux Group and East Coast South.

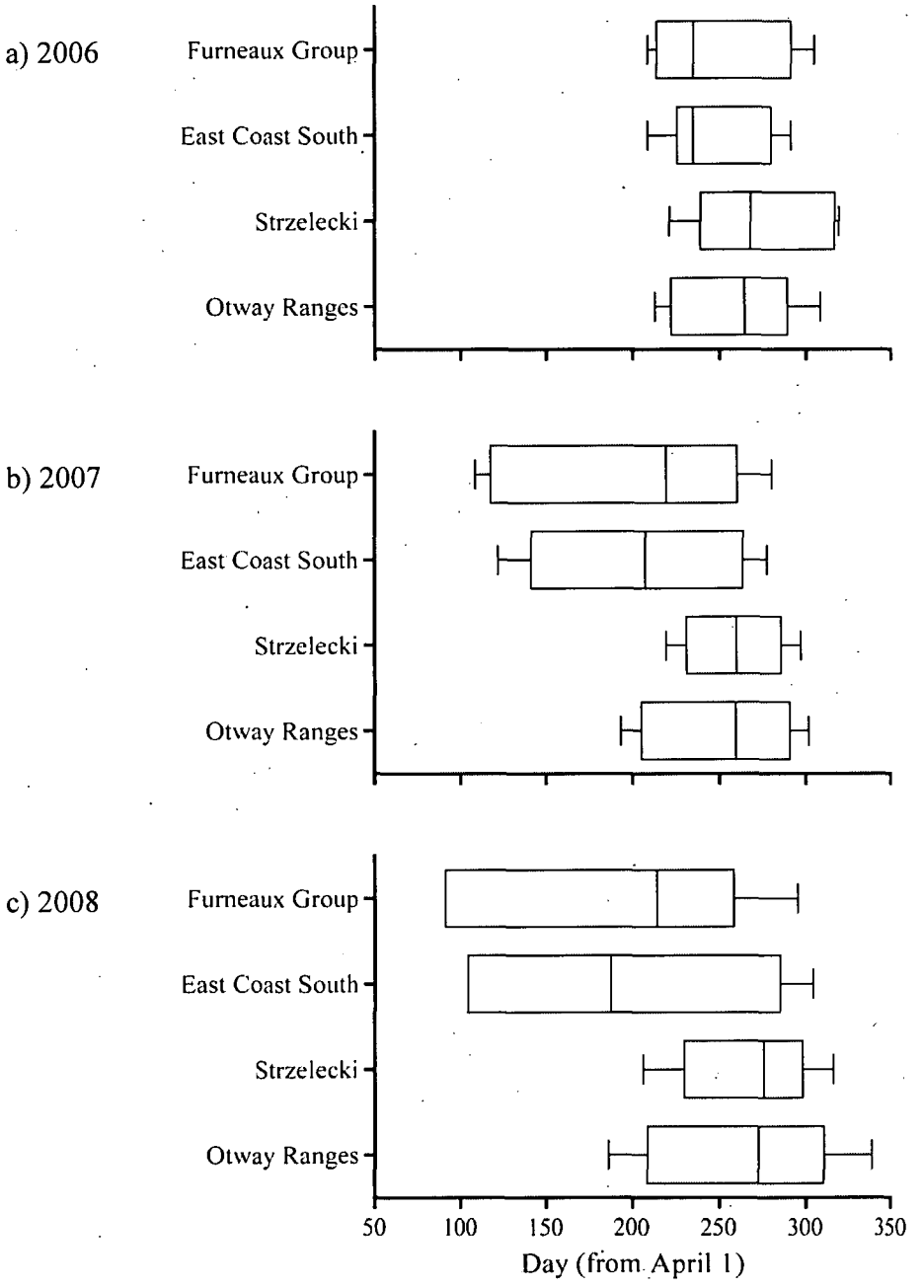


Figure 2.7. Overall and peak flowering period for each race of *globulus* each year at the *globulus* seedEnergy Cambridge orchard: a) 2006; b) 2007; c) 2008. Box shows the period of calculated peak flower opening for each race: minimum and maximum FTpeak, with mean FTpeak indicated by the line within the box; lines outside of the box indicate the period of flower opening, from observed minimum FT1 to maximum endfl.

Flower bud opening time was highly heritable with a broad sense heritability (H^2) of 0.69-0.77 depending on the flower opening measure used (Table 2.3). The majority of the phenotypic variation in the timing of flower bud opening was attributable to genetic variation i.e. between races and between genotypes within races (Figure 2.8). The genetic variation explained by differences among races (37.9%) and among genotypes within races (37.3%) were similar (Figure 2.8). Year was one of the largest environmental effects (5.0%), and year x genotype/race interactions were small but significant (Figure 2.8).

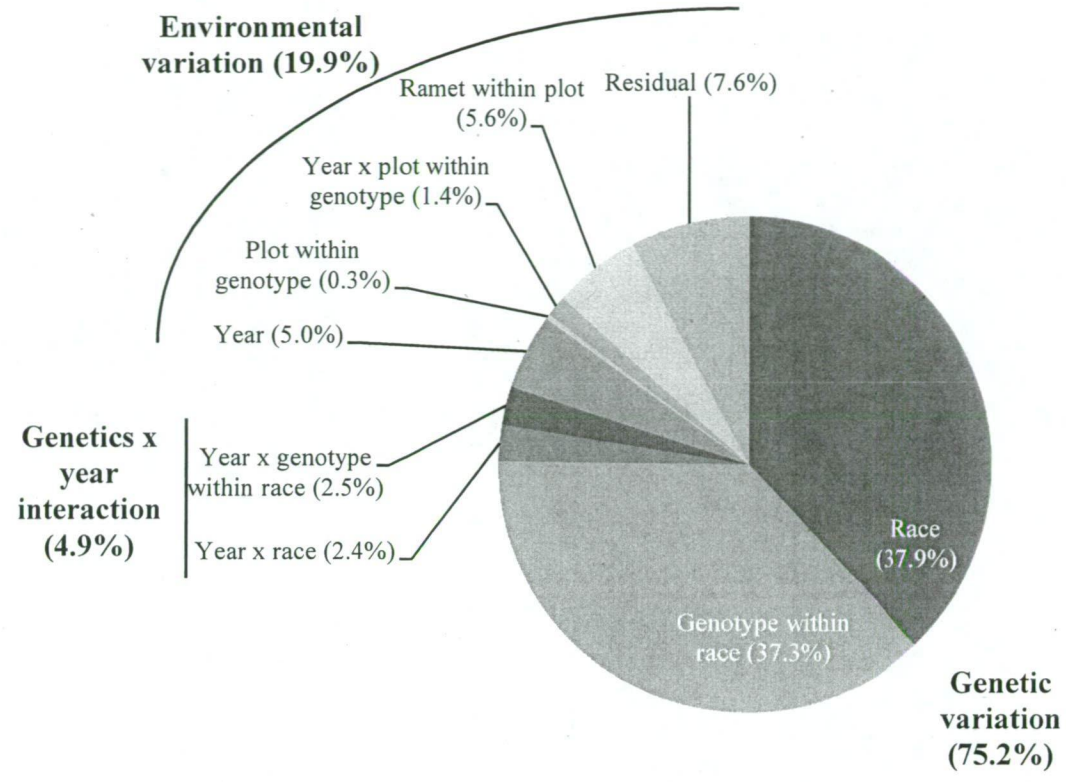


Figure 2.8. Partition of sources of phenotypic variation in FTpeak at the *globulus* seedEnergy Cambridge orchard, showing the percentage of variation that is attributable to genetic variation (among races and genotypes within races), interactive effects with year (year x genotype and year x race) and environmental variation (remaining components).

2.3.3 Flower bud development

As flower bud initiation occurs virtually simultaneously in all clones (see 2.3.1), but there is significant genetic variation in flower opening time (see 2.3.2), the difference in flower opening time is therefore mainly due to slower bud development in late flower opening genotypes. To investigate whether this difference was due to overall slower rates of bud development in the late flower opening genotype, or whether buds develop at the same rate but late flower opening genotypes sit dormant until flower opening time, bud growth was monitored in early (genotype 4489, Furneaux Group) and late (genotype 5296, Strzelecki Ranges) clones. Buds were initiated on both 4489 and 5296 trees in spring, between day 207 and day 221, but flower bud development (macroscopic initiation until flower opening) lasted 181 days in the 4489 ramet, and 376-378 days in the 5296 ramets. The buds of both genotypes grew at a similar rate until the end of summer (Figure 2.9). After this point, the buds of 4489 continued to grow and opened in late autumn, but the 5296 buds entered a period of dormancy which lasted until the following spring, after which there was a final growth spurt before opening (Figure 2.9). The outer (sepaline) operculum was also shed earlier in development in 4489 buds than in 5296 buds (Figure 2.9).

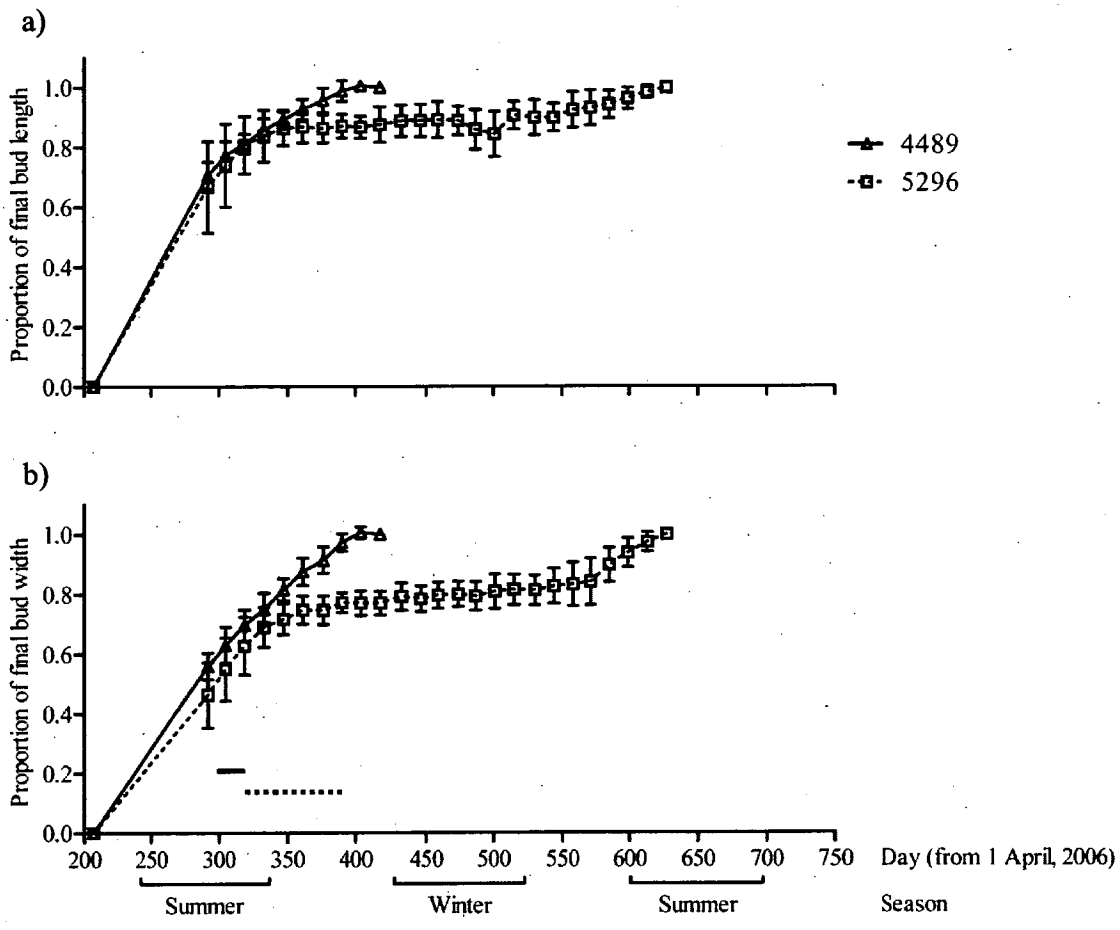


Figure 2.9. Mean bud a) length and b) width of an early (4489) and late (5296) flower opening genotype of *globulus* at the seedEnergy Cambridge orchard. Error bars show standard deviation. Means are based on a sample size of 21 buds for genotype 4489 and 32 buds for 5296. The period of shedding of the outer (sepaline) operculum for each genotype is indicated on the bottom graph (solid line, 4489; dashed line, 5296). Note there were no measurements taken between initiation and day 291.

2.3.4 *Association between climatic variables and flower bud development*

All of the daily climatic variables were positively correlated with each other (Table 2.8). The heat sum measures in particular were highly positively correlated at r=0.98 and above (Table 2.8).

Table 2.8. Pearson's correlations between daily climatic variables measured at a weather station close to the seedEnergy Cambridge *globulus* orchard. Significance levels are indicated (*** P < 0.001) and number of observations is given in parentheses.

	Solar	Sunshine	Heatsum4	Heatsum5
Sunshine	0.77*** (2259)			
Heatsum4	0.54*** (2269)	0.25*** (2349)		
Heatsum5	0.54*** (2269)	0.25*** (2349)	1.00*** (2359)	
Heatsum4min	0.56*** (2235)	0.26*** (2315)	0.98*** (2325)	0.98*** (2325)

On average, 2006 appeared to have a cooler late autumn to early winter than the other years, while 2005 had a warmer autumn than other years (Figure 2.10).

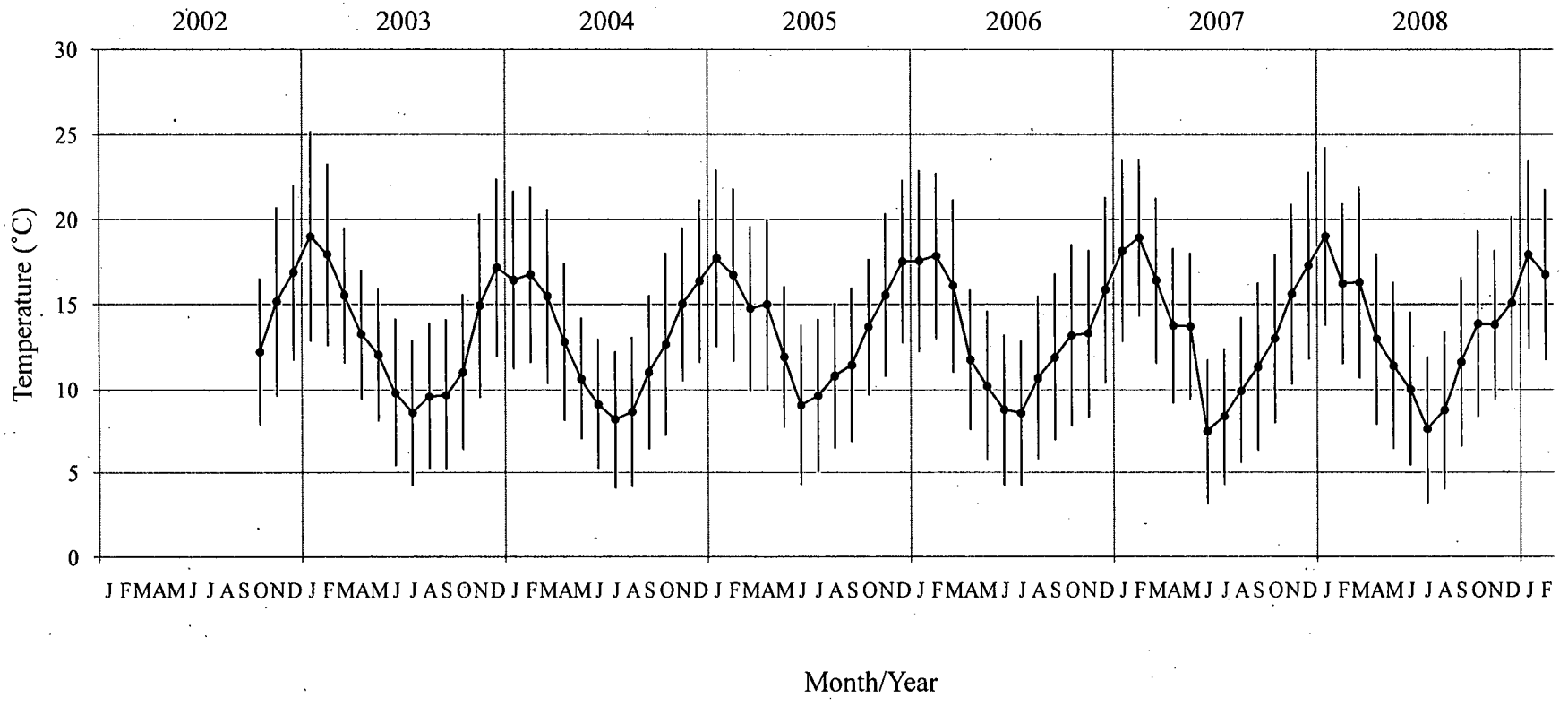


Figure 2.10. Average minimum, maximum and mean temperatures for the Hobart Airport weather station, which is only 6.5 km from the seedEnergy Cambridge *globulus* orchard and at similar elevation.

In the subset of samples for which there was a complete five year series of FT1 or FTpeak records (n=79 or n=70 respectively), there was a significant difference across years in accumulation of nearly all climatic variables during the ten different measures of FI to FT (Table 2.9). However there was no significant difference in mean Heatsum5 from FIestrace to FTpeak among years (Table 2.9, Figure 2.11a). This is despite there being a difference in mean FTpeak among most years for the same subset of trees ($F_{4,69}$ = 132, P < 0.0001, Figure 2.11b). For example, while these trees flowered later in 2006 than 2007 and 2008 (Figure 2.11b) the average accumulated heat sum (above a base temperature of 5°C) from FIobsrace until FTpeak was not significantly different in these three years, as the same heat sum was accumulated earlier in the year in 2007 and 2008 than in 2006 (Figure 2.11a), consistent with the cooler autumn in 2006 (Figure 2.10). In 2006 this also corresponded to delayed flowering of all races and a reduced flowering period, compared with 2007 and 2008 where there was greater flowering asynchrony among races and the flowering period of the orchard was longer (Figure 2.7).

Table 2.9. *F* values and significance of the year effect on accumulation of five climatic variables from FI to FT in the seedEnergy Cambridge *globulus* orchard. Five different FI measures and two FT measures were used, resulting in ten different time periods being tested. Number of significant (P<0.05) pairwise comparisons among years (out of ten pairwise comparisons when using FIest related measures of FI, or out of one pairwise comparison when using FIobs related measures of FI) is indicated in parentheses. The only combination for which there was no significant difference among years is shaded. *** P < 0.001; ** 0.001<P<0.01; * 0.01<P<0.05; ns not significant.

FI measure	FT measure	DF	Solar	Sunshine	Heatsum4	Heatsum5	Heatsum4min
FIobs	FT1	1,78	114.3*** (1)	70.3*** (1)	38.1*** (1)	38.6*** (1)	20.9*** (1)
FIobs	FTpeak	1,69	99.8*** (1)	65.0*** (1)	35.8*** (1)	36.0*** (1)	21.1*** (1)
FIobsMon	FT1	1,78	142.5*** (1)	53.3*** (1)	24.7*** (1)	23.4*** (1)	9.5** (1)
FIobsMon	FTpeak	1,69	129.5*** (1)	56.9*** (1)	27.8*** (1)	31.6*** (1)	13.7*** (1)
FIest	FT1	4,78	364.4*** (9)	135.1*** (8)	9.5*** (7)	8.9*** (5)	16.3*** (6)
FIest	FTpeak	4,69	618.9*** (10)	209.5*** (10)	6.0*** (2)	2.7* (1)	7.6*** (3)
FIestMon	FT1	4,78	417.6*** (9)	159.0*** (8)	36.3*** (7)	44.3*** (8)	47.6*** (8)
FIestMon	FTpeak	4,69	669.6*** (10)	242.1*** (9)	24.7*** (6)	29.4*** (7)	34.4*** (6)
FIestrace	FT1	4,78	330.9*** (9)	123.4*** (8)	10.4*** (6)	10.1*** (5)	19.2*** (6)
FIestrace	FTpeak	4,69	581.1*** (9)	187.7*** (9)	4.5** (2)	1.8ns (0)	8.8*** (3)

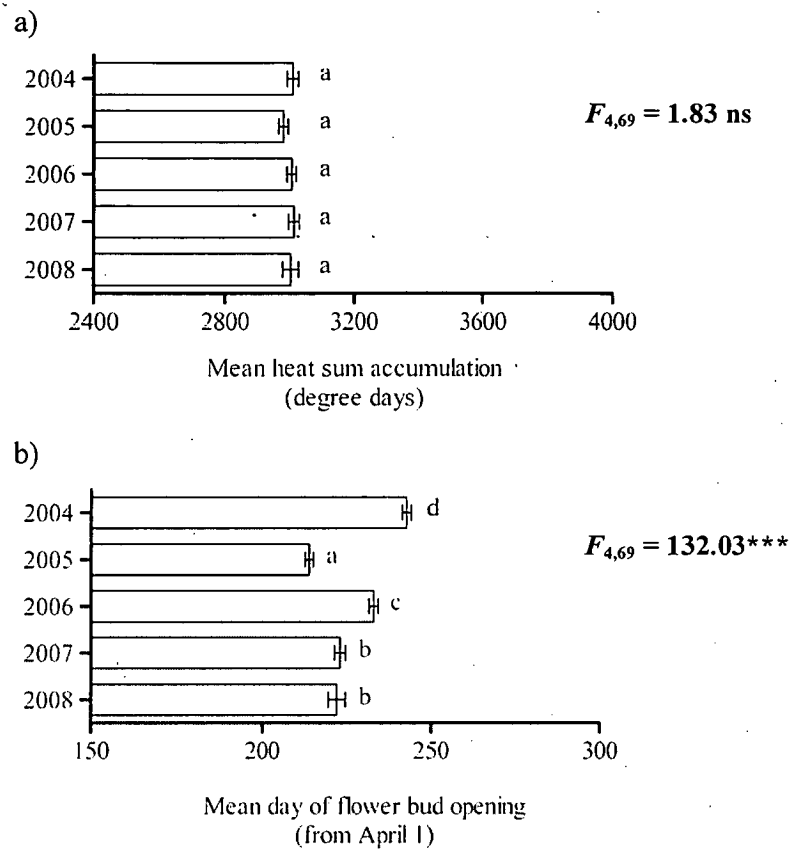


Figure 2.11. Mean a) Heatsum5 from FIestrace to FTpeak, and b) FTpeak each year in the 70 trees in the seedEnergy Cambridge *globulus* orchard that had five complete years of FTpeak data. Error bars represent one standard error, and year means sharing a letter were not significantly different using the Tukey-Kramer test.

When the mean Heatsum5 over the FIobsrace to FTpeak period was compared among years for each race, there were significant differences among years in some races (Figure 2.12). However, exactly the same individuals or genotypes were not necessarily used each year to obtain the mean Heatsum5 from FIobsrace to FTpeak, and this could account for the variation.

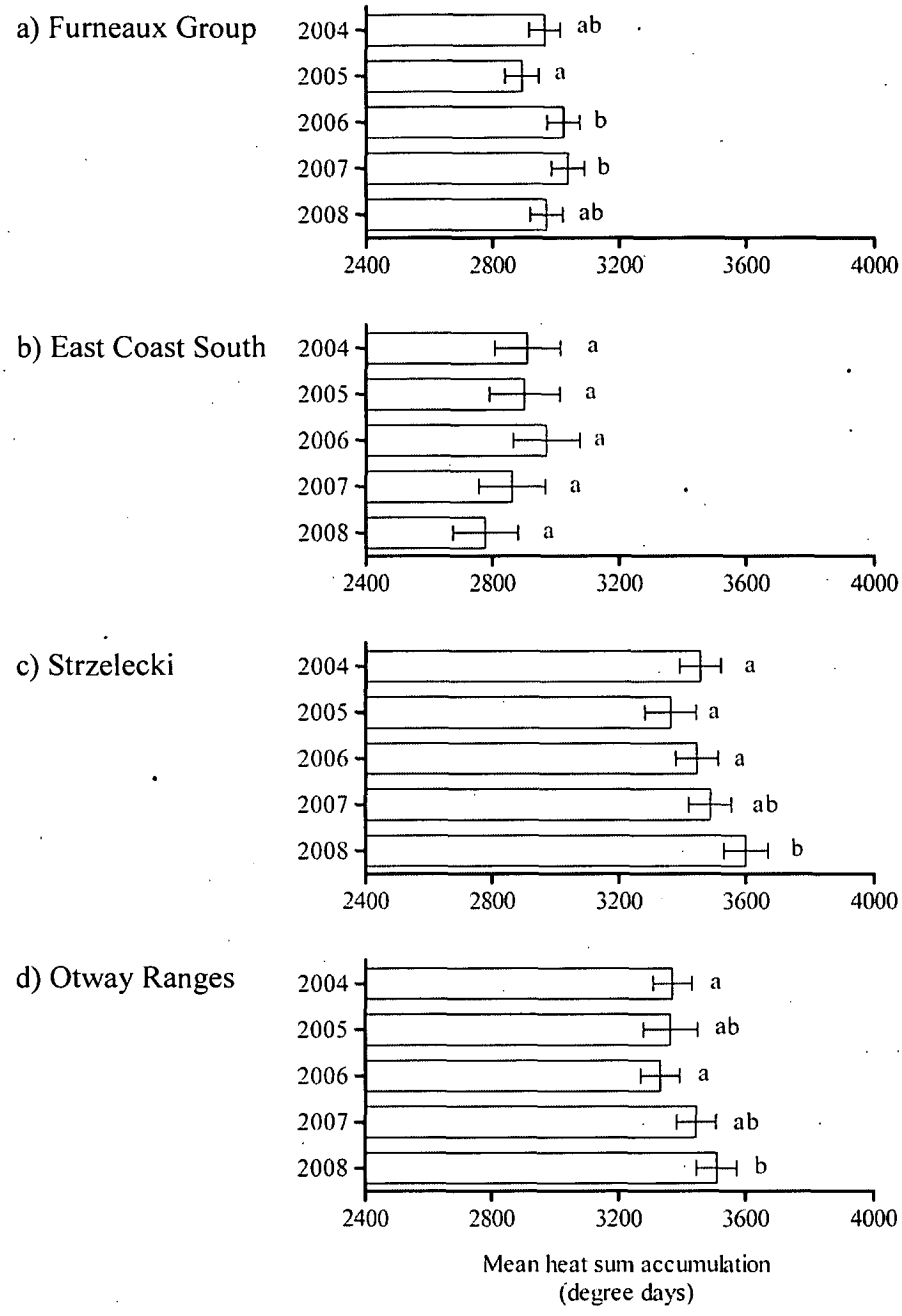


Figure 2.12. Least square means of Heatsum5 from FIestrace to FTpeak each year in each race at the seedEnergy Cambridge *globulus* orchard. Least square means were derived using the same repeated measures model used in Table 2.4 and Table 2.5. Error bars represent one standard error, and year means sharing a letter within a race were not significantly different using the Tukey-Kramer test. The number of individuals and representation of genotypes in each race was different each year (Table 2.1).

2.4 Discussion

2.4.1 Genetic variation in flowering traits

There was a large amount of variation in the timing and duration of flowering in *globulus*. Each year there was a significant difference among races in the timing of flower bud opening, with Furneaux Group and East Coast South races flowering earliest, and Strzelecki Ranges and Otway Ranges races flowering latest, which, to some extent, matches the natural pattern of variation observed in 2006 (Potts and Vaillancourt, unpubl. data, Figure 2.1). There was, however, a significant year effect, which indicates that flower opening time is not a photoperiod-mediated response in *globulus*, though photoperiod could still play a part in the onset of flower opening. Though the onset of flower opening in the orchard varied among years, the rank order of flower opening at the tree/genotype level was similar among years (i.e. Spearman's rank correlation was high). However, in some years (e.g. 2006) the overlap in flowering period between the early and late races was greater than in other years (e.g. 2007 and 2008), so the barrier to gene flow between the early and late flowering races is unstable to some degree. This has implications for management of the seed orchard, as synchronous flowering is important to ensure efficiency of OP seed production. Non-synchronous flowering is also a major problem in conifer seed orchards though 30-day flowering period for a conifer seed orchard is considered asynchronous (Nikkanen 2001), a much shorter flowering period than is observed in the Cambridge *globulus* orchard. Conversely, in *globulus* plantations, asynchronous flowering of provenance and native stands would be desirable to minimise the risk of genetic pollution. In more recent years, germplasm from *globulus* provenances from mainland Australia (Otway Ranges and Strzelecki Ranges) or Furneaux has been established in Tasmania. Mainland Australian provenances are differentiated from Tasmanian provenances (see Chapter 1) but there are clearly pre-mating barriers to gene flow between these provenances, reducing the risk of gene flow. The strength of this pre-mating barrier to gene flow does, however, vary from year to year. By contrast, there does not appear to be a pre-mating barrier to gene flow between the Furneaux and Southern Tasmania races but these two provenances are closely related (see Chapter 1), which reduces the potential impact of gene flow.

While the flowering period of the orchard occurred over a period of up to eight months (in 2008) and there was a large amount of genetically-based variation in flower opening time, flower bud initiation occurred nearly simultaneously in all clones each year. Late flowering genotypes, if anything, initiated flower buds slightly earlier than other genotypes. Therefore, the differences in flower opening time were mainly due to a longer period of flower bud development in late flowering genotypes than in early flowering genotypes (see 2.4.3). There was a significant difference in mean FIobs between the two years, but this may have been an artefact of scoring inexperience in detecting new flower buds in the first year of FIobs surveys, as the mode day of FIobs was the same in both years. Further surveys are needed to determine if the day of flower bud initiation is the same each year, which could indicate it is a photoperiod-mediated process.

Other phenological studies of perennial systems have reported a correlation between leafing date or bud flush (equivalent to Veg in this study) and flowering date (e.g. sweet cherry, DeVries 1967; apple, Visser and Schaap 1967, Mehlenbacher and Voordeckers 1991; and pistachio, Chao and Parfitt 2003) but this correlation was not observed in this study. Rather, in 2007 the timing of vegetative flush was moderately positively correlated with observed flower bud initiation time, with the overall mean vegetative flush four days earlier than the overall mean flower bud initiation time.

2.4.2 Heritability of flowering traits

There was very little genetic based variation in flower bud initiation time, with most variation due to environmental effects (FIobs $H^2 = 0.06$). However, flower bud opening time was highly heritable in the broad sense (FTpeak $H^2 = 0.77$), with approximately equal variation attributed to race or genotype within race effects. Flowering time traits are often highly heritable, for example days to first flowering in annuals (e.g. cowpea $H^2 = 0.95$, Machado *et al.* 2008; and lentil $H^2 = 0.94$, Bicer and Sakar 2008) and seasonal flowering in perennial species [e.g. days to peak flowering in almond narrow sense heritability (h^2) = 0.99, Dicenta *et al.* 1993; days to receptivity in black pine $H^2 = 0.70$, Matziris 1994; days to peak flowering in pistachio $h^2 = 0.79$-0.89, based on half-sib and parent-offspring regression analysis respectively, Chao and Parfitt 2003;

and days to peak flowering in kiwifruit $h^2 > 0.85$, Cheng *et al.* 2006]. A previous study of the genetic control of flowering time in *globulus* calculated similarly high heritabilities of flowering time traits (equivalent of FT1 $h^2 = 0.81$, equivalent of FTpeak $h^2 = 0.65$, Gore and Potts 1995) and intermediate inheritance of racial differences in flowering time in inter-race hybrids, suggesting most of the variation is due to additive genetic effects (Gore and Potts 1995). In contrast to flower opening time, flowering duration in perennials is under weak or variable genetic control (Dicenta *et al.* 1993; Cheng *et al.* 2006), a pattern that was also observed in this study (flperiod $H^2 = 0.09$). In a previous *globulus* study though, narrow-sense heritability of flowering period was moderate ($h^2 = 0.35$, Gore and Potts 1995). This may be partly because the flowering period is positively correlated with flower abundance, which could be under moderate genetic control.

The high heritability and intermediate inheritance of flower opening time implies that, if desired, careful crossing and selection could produce cultivars for seed orchards that have less variation in flower opening time, which could reduce the flowering asynchrony observed in some years in seed orchards such as at Cambridge. This could not only increase the pollination efficiency in the orchard, but also reduce labour costs in harvesting seed, as the seed is usually picked from the trees twelve months after the commencement of flower opening to allow sufficient time for seed development and maturation (P.L. Gore pers. comm.) and in years when the overall flowering period of the orchard is long, the harvest period the following year is therefore also extended. Harvesting fruit as early in the season as possible is desirable, to allow enough time for cleaning and testing germination of seed prior to dispatch to nurseries. Current work at UTAS aims to address this problem by investigating the effects of early fruit harvest on germination (Rix *et al.* in prep).

2.4.3 Variation in dormancy among early and late flowering clones of globulus

The later flower opening time of genotype 5296 was not entirely due to a slower rate of bud development *per se*, but instead was mainly due to an inactive period of around six months over winter where the flower buds were dormant, followed by a final growth spurt in spring, leading to flower bud opening. The timing of release from dormancy

may therefore be the direct driver of flower opening time. In almond, flowering is a two stage process, with a winter dormancy phase that is terminated once the accumulation of exposure to low temperatures is achieved (chilling requirement) and a second stage in which buds develop at a rate depending on temperature (heat sum requirement), and the chilling and heat sum requirements may differ between cultivars (Rattigan and Hill 1986). This model may also account for the variation in flower bud development rates among genotypes of *globulus*. It may be that there is no chilling requirement in early genotypes such as 4489, and so the buds do not undergo a dormant period, while in later genotypes a chilling period is required, and the heat sum after this chilling period determines flower opening time.

The differences in bud development time may also be due to the differences in bud morphology. Bud morphology appears to have a strong genetic basis (pers. obs.) though this has not been tested statistically. Pure *globulus* (East Coast South and Furneaux) has large, solitary buds, while *globulus* intergrades (Otway Ranges and Strzelecki Ranges) may have up to three buds per umbel, and the buds are often smaller (see Chapter 1). The early genotype (from Furneaux Group) monitored in this study had large solitary buds while the late genotype (from Strzelecki Ranges) had smaller buds in umbels of three. The earlier flowering time of the genotype with larger buds is in agreement with Griffin (1982), who found that eucalypt species with larger flowers tended to flower earlier than those with small flowers. Primack (1987) hypothesised that plant species with larger buds (and therefore larger fruit) flower earlier in the year as the larger fruits take longer to mature. In an analysis of 51 eucalypt species, Keatley and Hudson (1998) gave further evidence to support this hypothesis, indicating there was a link between fruit or umbel volume and flowering time in eucalypts. Early flowering was a characteristic of species with large buds and fruits and those with a large overall umbel volume. The use of just one genotype each of an early and late flower opening genotype is insufficient to lend support to Primack's (1987) hypothesis, but in future studies it would be interesting to test the Spearman rank correlation between bud size, flower opening time and fruit ripening time in *globulus*, especially in a random mating population segregating for all these variables.

2.4.4 *Role of climate in the onset of flower opening*

While the variation in flower opening time is under strong genetic control, there was still a significant environmental component, and year was one of the largest of the environmental effects. There was a precedent for heat sum causing the year-to-year or site-to-site variation in annual flower opening time in eucalypts (see 2.1.4). Indeed, when comparing trees for which there was a complete five year flowering record, the mean heat sum (above 5°C, "Heatsum5") of these trees explained the variation in mean timing of flower opening for these trees better that the day of the year. The between-year variation in mean flowering time of these selected trees was up to 30 days and highly significant ($F_{4,69}$ = 132.03***), but when defined on the basis of heat sum, the variation was smaller and not significant ($F_{4,69}$ = 1.83 ns). This provides good evidence that, overall, heat sum is a major driver of the year-to-year variation in flower opening time observed in this orchard. Indeed, the later flowering of 2006 can be attributed to the cooler autumn of that year.

When the heat sum data was compared for all trees (i.e. rather than just the selected trees for which there was a complete five year flowering record), there was some variation in the heat sum accumulated per race among years. This could be because different genotypes were represented each year, and these genotypes within the races have different heat sum requirements. In addition, the model was based on one climatic variable at a time, and there could be interactions between these variables. For example, heating from solar radiation and cooling from winds can cause a discrepancy between air and bud temperatures. Also, other studies have shown that, in addition to heat sum, variation in rainfall can affect phenological development in some trees while others are not affected (Spano *et al.* 1999). However rainfall could not be incorporated into the model in this study as the rainfall at the orchard is supplemented by irrigation.

Many models of phenological development in trees incorporate chill units as well as heat sum, where a certain chilling requirement is necessary to break bud dormancy, followed by a heat sum requirement for flower development to anthesis. In almond, a combined chill unit / heat sum model based on data from one location could be used to predict flowering time at a different location with a different climate (Rattigan and Hill 1988). The pattern of flower bud growth in the early (4489) and late (5296) flower opening

genotypes, in which a constant bud growth rate was observed in genotype 4489, but the buds of genotype 5296 underwent a dormant period, could indicate that the combination of chilling and heat sum requirement may differ from genotype to genotype, and the early flower opening genotypes may have no chilling requirement at all, as the buds do not have a period of dormancy. This is worth investigating in future studies, but would require the bud development to be measured in more genotypes in order to determine the timing of dormancy break.

It has been shown that there is probably a specific heat sum associated with each race, but these are not necessarily heat sum requirements *per se*. It would be necessary to investigate the flower opening time of the same genotypes planted at other sites with different climates to elucidate whether the accumulated heat sum measured for each genotype at the Cambridge orchard is a specific heat sum requirement. The genotypes planted at the Cambridge orchard are part of the STBA breeding population of *globulus*, and the same genotypes have also been planted in seed orchards in Western Australia and South Australia. In this study, MSP records were consistent with flowering surveys and therefore formed part of the 2008 flower opening time data. MSP records from the other STBA seed orchards could be a cost-effective means of obtaining a large data set of flower opening time of these same genotypes growing in a range of climates.

There is some indication that the early flowering races are perhaps more sensitive to climate than the later flowering races, because in warmer years, the onset of flower opening in the early flower opening races advanced more than in the late flower opening races. Also, the flowering period of early races was longer in the warmer years, whereas the flowering period of late races was similar between years, and this resulted in a greater flowering asynchrony in the orchard in warmer years. This has obvious implications for future management of *globulus* seed orchards as there is clear evidence of global warming, including within the *E. globulus* range where an annual mean temperature increase of 2-3°C is predicted (Christensen *et al.* 2007). In natural stands of *globulus*, the synchrony of *globulus* flowering and swift parrot migration could be disrupted if the increase in temperature affects the phenology of the two species differently. This is the case in the winter moth egg hatch-oak bud burst system where the moth egg hatch date has advanced more than the oak bud burst date over the last

twenty years (van Asch *et al.* 2007). The high levels of genetic variation in flower opening could, however, give *globulus* the evolutionary flexibility to respond to future shifts in temperature and pollinator activity.

This study has shown that there is a large amount of genetically based variation in flower opening time in *globulus*, with heat sum a major part of the environmental effect. The next step will be to identify the genes that may be controlling the variation in these traits, from a list of candidate genes that are known to be associated with flowering processes in the model annual plant *Arabidopsis*, or other forest trees (Chapter 3). This will eventually contribute towards a predictive model of the flowering time of plantations and seed orchards, based on knowledge of the genetic material deployed and the climate of the local area.

3 SEASONAL VARIATION IN EXPRESSION OF FLOWERING GENES IN AN EARLY AND A LATE FLOWER OPENING GENOTYPE OF *E. GLOBULUS* SSP. *GLOBULUS*

3.1 Introduction

3.1.1 Life cycles of plants

3.1.1.1 Annual plants

Annual plants are characterised by a short, simple life cycle consisting of a vegetative phase and a single reproductive phase. After flowering and producing fruit and seeds, the plant dies. The onset of the reproductive phase is influenced by environmental cues such as photoperiod, temperature and light quality, and internal signals such as hormones (Boss *et al.* 2004). In most annual plants, flower initiation is followed directly by flower development and opening, and the term "flowering" refers to these two processes in combination.

Photoperiod is sensed in the leaves, and long-day and short-day plants flower in response to short and long dark periods respectively. *Arabidopsis thaliana*, a small, weedy annual in the family Brassicaceae and a model species for plant biology, is a facultative long-day plant. Flowering is promoted by decreasing night length, which predicts the onset of spring and summer, and flowering is delayed in short-day conditions. On the other hand, most cultivars of rice are short day, and flower more rapidly in short-day conditions than in long-days. In plants of tropical origin, such as rice, this response means that flowering and seed set coincides with the onset of the rainy season (Battey 2000). *Arabidopsis* plants originating from cool temperate regions are winter annuals; they germinate in the summer, grow vegetatively over winter, and flower under the long days of spring, but only if exposed to a period of cool temperatures, often referred to as vernalisation. This is in contrast to summer annuals that germinate and flower in the same summer, without the vernalisation requirement (Mouradov *et al.* 2002).

In *Arabidopsis*, flowering is also promoted by gibberellic acid (GA, reviewed in Mouradov *et al.* 2002) and overcrowding, as detected by changes in light quality (Putterill *et al.* 2004).

3.1.1.2 Perennial plants

Plants that have the ability to live and reproduce for two or more years are termed perennial plants. Unlike annual plants, the shoots of perennial plants are exposed to repeated seasonal environmental changes. Many perennial species cease growth before the cold or dry season, with their vegetative buds remaining dormant until favourable environmental conditions return. This response is promoted by short days in many woody species (Wareing and Saunders 1971). Another adaptation for survival by perennial plants is the deciduous habit, where plants abscise their leaves during cold or dry seasons. In woody perennials the trunk and branches will persist, whereas the above-ground parts of some herbaceous perennials may die back entirely in the cold or dry season, returning in spring (e.g. *Narcissus*). In trees, temperature has a key role in the timing of vegetative bud burst: most species require a period of vernalisation to break dormancy, and once dormancy is broken, the development of the vegetative buds to bud-burst is dependent on the accumulation of heat (Battey 2000).

Woody perennials are also characterised by a juvenile period that can last years to decades. There are multiple tissues and organs that exhibit phase change, and these are regulated independently (reviewed in Poethig 2003). The most obvious phase change is the transition to reproductive development, and once reproductively mature, trees can alternate between the production of vegetative and reproductive shoots, on a single branch. The shoot therefore goes through repeated phase change cycles between vegetative and reproductive growth, a phenomenon that is absent in annual plants such as *Arabidopsis* (Boss *et al.* 2004). In addition, flowering and non-flowering branches can exist on the same tree, and reproductively competent trees do not flower every year, additional characteristics that do not have counterparts in *Arabidopsis*.

Flower initiation in woody perennials is influenced by environmental and internal signals, but these cues may result in a contrasting response to that in annual plants. Seasonal flower initiation in woody perennials tends to be less influenced by

photoperiod; instead, flower initiation is determined by alternative environmental signals such as vernalisation (e.g. mango), or may be autonomous (e.g. apple) (Wilkie *et al.* 2008). Also, in contrast to many annual plants, GA inhibits rather than promotes flowering in many woody angiosperms and the GA-inhibitor paclobutrazol is often used to stimulate flowering in trees of commercial importance (Meilan 1997). Conversely, in many gymnosperm tree species, flowering is induced by GA application (e.g. in the Pinaceae, Pharis *et al.* 1987).

In many perennial species the delay between flower initiation and flower opening is longer than in annual plants. This may be the reason that perennial species are less responsive to photoperiod than annual plants, as flower initiation in perennials bears less relation to flower opening time and fruiting (Battey 2000). For example, many tree fruit crops, particularly those of temperate origin, have a marked delay between flower initiation and opening, and this is also the case in herbaceous perennials like strawberry (Battey 2000). Furthermore, temperate horticultural trees tend to exhibit a period of dormancy between flower initiation and flower opening, whereas in tropical species, flower development tends to be continuous (Wilkie *et al.* 2008).

3.1.1.3 *Eucalyptus globulus*

Like most forest trees, *Eucalyptus globulus* ssp. *globulus* (hereafter referred to as *globulus*) has a long juvenile, non-flowering phase followed by an annual cycling between vegetative and reproductive growth (see 2.1.1, Figure 2.2). The flower bud morphology of *Eucalyptus* is unusual and is described in 2.1.1 (Figure 2.3). The shedding of the inner (petaline) operculum to expose the stamens occurs once all parts of the flower bud are fully developed, from late autumn to early summer in *globulus* (see Chapter 2) and is the process commonly referred to as flowering. Here, the shedding of the inner operculum will be termed "flower opening" to avoid confusion with the flower bud initiation stage, as the latter corresponds to the stage normally referred to as "flowering" in *Arabidopsis*. In *globulus*, the timing of flower opening has a strong genetic basis and is also influenced by climate rather than photoperiod, while the timing of flower initiation is under weak genetic control (Chapter 2).

Like *Arabidopsis*, flower initiation in *Eucalyptus* is affected by vernalisation, overcrowding and gibberellins, though the response in eucalypts is not always the same as in *Arabidopsis*. Vernalisation promotes flower bud initiation in some eucalypt species (Moncur 1992, Moncur and Hasan 1994) including *globulus* (Hasan and Reid 1995). However in contrast to *Arabidopsis*, eucalypt flower bud initiation appears to be reduced by overcrowding. Trees that were widely spaced in trials of *E. nitens* produced significantly more flowers than closely spaced trees (Moncur *et al.* 1994, Williams *et al.* 2006) and flower abundance was lower on trees in inside rows than on trees in outside rows of *globulus* plantations (Barbour *et al.* 2008b). Application of GA inhibitors increases the intensity, but does not affect the timing, of flower bud initiation (Hetherington *et al.* 1992, Griffin *et al.* 1993, Hasan and Reid 1995) and application of the GA inhibitor paclobutrazol has become standard management practice to increase flowering intensity and reduce internode length in *globulus* seed orchards. However, in the absence of winter cold, the chemical is relatively ineffective (Moncur and Hasan 1994) indicating that vernalisation is the most important process, at least in *E. nitens*. There is also some evidence that water stress can promote floral induction in *globulus* (Hasan and Reid 1995).

3.1.2 Genetic pathways to flowering

3.1.2.1 The flowering pathway in *Arabidopsis*

The transition to flowering in plants is influenced by external signals such as light and temperature, and internal signals, which interact through a complex genetic pathway mainly studied in the annual plant, *Arabidopsis* (Figure 3.1, reviewed in Boss *et al.* 2004). *Arabidopsis* is an excellent model species for flowering biology for a number of reasons: it has a relatively small genome that has been sequenced and is publicly available; its small size and simple growth requirements make it easy to grow in the laboratory; it is easily transformed (i.e. genetically modified) and it has a short life cycle of around eight weeks (Somerville and Koornneef 2002). In addition, it has a large amount of genetic variation across its widespread natural distribution, and this can be exploited to map and isolate genes of interest.

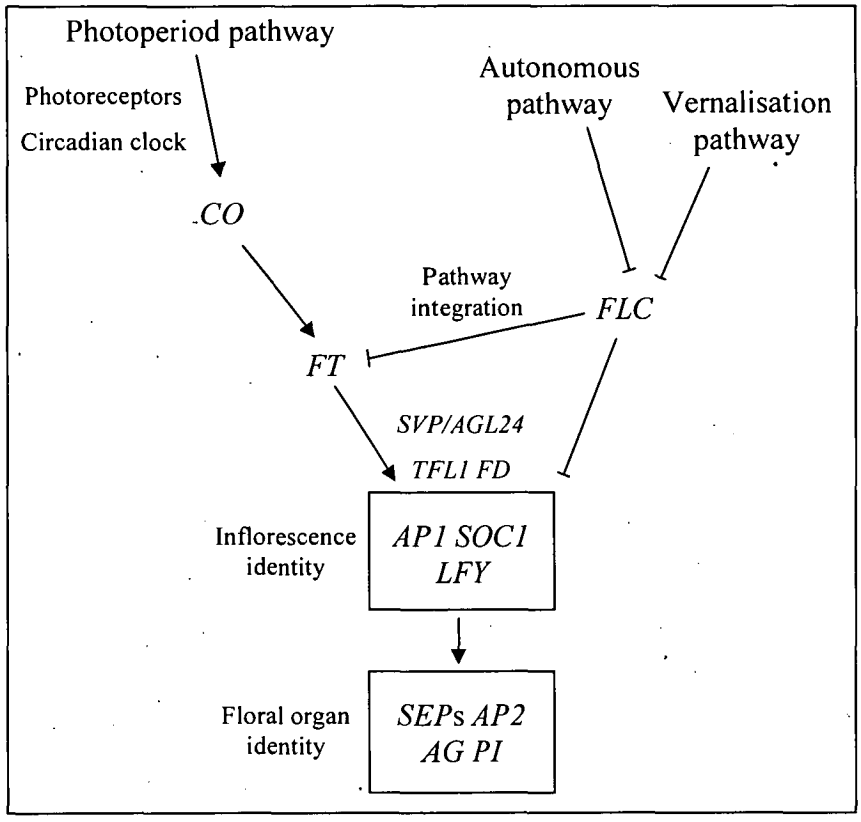


Figure 3.1. Simplified diagram of the genes in the flowering pathway of the annual plant *Arabidopsis thaliana*. Mallet symbols indicate inhibition; arrows indicate promotion.

In *Arabidopsis*, the photoperiod, vernalisation and endogenous pathways converge at a common set of genes to enable the integration of these different signals into a coordinated flowering response (Figure 3.1). *FLOWERING LOCUS T* (*FT*) is expressed in the leaf and the FT protein moves to the apex (Corbesier *et al.* 2007), where it interacts with the transcription factor FD (Wigge *et al.* 2005) to activate inflorescence identity genes including *APETALA1* (*AP1*) and *LEAFY* (*LFY*), which trigger flowering (Putterill *et al.* 2004).

Most of the genes in the flowering pathway belong to three large gene families: MADS-box, *FT / TERMINAL FLOWER1* (*TFL1*), and *CONSTANS-LIKE*. MADS-box proteins are transcription factors that have a highly conserved DNA binding domain (the "MADS-box") and are key regulators of a range of plant development processes (Parenicova *et al.* 2003). Of the five clades of MADS-box genes, the genes of the MIKC clade have been studied in the most detail, and these include genes involved in the control of flowering time (*FLOWERING LOCUS C* [*FLC*], *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1* [*SOC1*], *SHORT VEGETATIVE PHASE* [*SVP*], *AGAMOUS-LIKE24* [*AGL24*], *APETALA1* [*AP1*]) and floral organ identity (*SEPALLATA* [*SEP1* to *SEP3*], *PISTILLATA* [*PI*] and *AGAMOUS* [*AG*]). The *FT / TFL1* gene family consists of six genes in *Arabidopsis*, including the mobile flowering stimulus *FT* and the flowering repressor *TFL1*. Other genes in this family are not as well-studied, but may also be involved in flowering. The *CONSTANS-LIKE* gene family consists of 17 genes in *Arabidopsis*, and includes the *CONSTANS* (*CO*) gene, which encodes a transcription factor that integrates circadian clock and light signals (Putterill *et al.* 2004), but little is known about the function of the other *CONSTANS-LIKE* genes (*COL1* to *COL16*).

The variation among plant species in their flowering response to environmental signals is important for adaptation to a wide variety of environmental conditions, to ensure that flowering occurs at the most favourable time for reproduction and dispersal. Although much has been learned from *Arabidopsis*, in order to understand the diversity of flowering responses, there is a need to study related flowering pathways in other plant species. The genes of the photoperiod pathway are conserved in short-day cultivars of rice, but the rice *CO* ortholog promotes flowering under short days instead of long days (Kojima *et al.* 2002, Hayama *et al.* 2003). On the other hand, the role of *FLC* is played by a different, unrelated, gene in wheat (*Triticum aestivum*, a long day plant, Yan *et al.* 2004) and *FLC* appears to be absent from four legume species (Hecht *et al.* 2005) and rice (Izawa *et al.* 2003).

The function of the downstream integrator genes appear to be well-conserved across many different plant groups (Putterill *et al.* 2004) but these sometimes play other roles in

plant development. For example, in pea, as well as having a role in flower initiation, *LFY* is involved in leaf development (Hofer *et al.* 1997).

3.1.2.2 Flowering genes isolated in trees

In the case of trees, characterised by a juvenile, non-flowering phase lasting years, and an annual cycling between vegetative and reproductive growth, the regulation of flowering time is potentially more complex than in *Arabidopsis* and other annual species. Homologues of some of the genes in the *Arabidopsis* flowering pathway have, however, been identified in many trees (Table 3.1), and in most cases their function has been supported by transgenic complementation of *Arabidopsis* mutants. The presence of these genes in a wide range of plant groups, including non-angiosperms such as *Ginkgo* and the gymnosperms, suggests that the regulation of reproductive initiation is evolutionarily conserved.

Table 3.1. Homologues of some of the *Arabidopsis* flowering pathway genes (listed in Figure 3.1) that have been isolated in key tree genera (excluding *Populus* and *Eucalyptus*, which are discussed in the text).

Gene	Genus[1]	Reference	Genbank accession number(s)
CO	*Prunus*	Silva *et al.* 2005	BU042239
FT	*Citrus*	Kobayashi *et al.* 1999	AB027456
	Malus	Haettasch *et al.* unpubl.	DQ535887
	Picea	Gyllenstrand *et al.* 2007	EF601975
	Prunus	Silva *et al.* 2005	BU044758
SVP	*Prunus*	Yamane *et al.* 2008	AB437345
TFL1	*Citrus*	Pillitteri *et al.* 2004a	AY344254
	Malus	Kotoda and Wada 2005	AB052994
	Metrosideros	Sreekantan *et al.* 2004	AY170872, AY170873
	Picea	Holefors *et al.* unpubl.	EF633467
	Prunus	Silva *et al.* 2005	BU574411
AP1	*Betula*	Elo *et al.* 2001	X99653
	Citrus	Pillitteri *et al.* 2004b	AY338974, AY338975
	Malus	Yao *et al.* 1999	AJ000759
	Metrosideros	Sreekantan *et al.* 2004	AY170871
	Prunus	Silva *et al.* 2005	BU039475
	Salix	Fernando and Zhang 2006	DQ068268
SOC1	*Citrus*	Tan and Swain 2007	EU032531,EU032532
	Malus	Mahna *et al.* 2006	DQ846833
LFY	*Cedrela*	Dornelas and Rodriguez 2006	AY633621
	Citrus	Pillitteri *et al.* 2004b	AY338976
	Malus	Wada *et al.* 2002	AB056158, AB056159
	Metrosideros	Sreekantan *et al.* 2004	AF007869
	Picea	Carlsbecker *et al.* 2004	AY701763
	Pinus	Mouradov *et al.* 1998	PRU76757
	Prunus	Silva *et al.* 2005	AY947465
	Hevea	Dornelas and Rodriguez 2005b	AY639378
SEP	*Betula*	Lemmetyinen *et al.* 2001	AJ252070
	Elaeis	Alwee *et al.* 2006	EL686357
	Prunus	Xu *et al.* 2008	BQ102369, EF440351, EF440352
AP2	*Larix*	Guillaumot *et al.* 2008	DQ988340, DQ988341
	Malus	Zhou *et al.* 2006	DQ083000
	Picea	Vahala *et al.* 2001	AF253970, AF253971
AG	*Betula*	Lemmetyinen *et al.* 2004	AJ252071
	Elaeis	Adam *et al.* 2006	AY739698, AY739699
	Ginkgo	Jager *et al.* 2003	AY114304
	Magnolia	Kim *et al.* 2005b	AY936228
	Malus	van der Linden *et al.* 2002	AJ251118
	Picea	Rutledge *et al.* 1998	PMU69482
	Prunus	Gao *et al.* unpubl.	EU072354
PI	*Betula*	Jarvinen *et al.* 2003	AJ488589
	Magnolia	Kim *et al.* 2005a	AY821779
	Malus	Yao *et al.* 2001	AJ291491
	Prunus	Zhang *et al.* 2008	EU005663

[1] Genera searched: *Abies, Acer, Betula, Castanea, Cedrela, Citrus, Elaeis, Fagus, Ficus, Fraxinus, Ginkgo, Hevea, Larix, Lithocarpus, Magnolia, Malus, Metrosideros, Picea, Pinus, Prunus, Quercus, Salix.*

As poplar grows rapidly, is easily transformed, and its genome sequence is now publicly available, data on the control of flowering in this species is accumulating. Comparison of the *Arabidopsis* and poplar sequences has shown that homologues of known *Arabidopsis* genes in the photoperiod, autonomous, vernalisation and GA signaling pathways, are nearly all present in poplar, as are the integrators of these pathways and the repressor genes not associated with any particular pathway (Brunner and Nilsson 2004). One exception is the floral repressor *FLC*. Homologues of the genes that are upstream of *FLC* in the *Arabidopsis* flowering pathway are, however, present in poplar, suggesting that the role of *FLC* may be played by a different, unknown, gene (Brunner and Nilsson 2004).

Transgenic expression of flowering genes in trees have shown that some genes, at least later in the flowering pathway, are functionally conserved, and can be used to promote early flowering in trees. *Arabidopsis LFY* induces early flowering when overexpressed in transgenic poplar and citrus trees (Weigel and Nilsson 1995, Rottmann *et al.* 2000, Pena *et al.* 2001), though in poplar the flowers were abnormal. Transgenic expression of *Arabidopsis AP1* also induces flowering in citrus (Pena *et al.* 2001). Transformed apple plants expressing the antisense apple homologue of *TFL1* flowered precociously (Kotoda *et al.* 2006) and juvenile *Populus* stems transformed with *Populus* homologues of *FT* initiated normal inflorescences within weeks to months of transformation, the first reports of juvenile transgenic trees producing inflorescences (Bohlenius *et al.* 2006, Hsu *et al.* 2006).

In *Populus*, *FT* is involved in regulating the age of first flowering, with trees exhibiting a gradual increase in levels of *FT* until the time of first flowering (at age 5-6, Bohlenius *et al.* 2006; or at age 11, Hsu *et al.* 2006). The authors postulated that epigenetic mechanisms may play a role in the regulation of *FT* in trees, and that each annual cycle of growth in juvenile trees leads to a gradual release of the chromatin-based repression. Modification of chromatin structure plays a key role in repressing floral development during embryo and vegetative development in *Arabidopsis* and similar processes may maintain the juvenile phase in trees.

Expression experiments in *Populus* have suggested that *FT* is also associated with the initiation of seasonal flowers (Hsu *et al.* 2006) and, surprisingly, the photoperiodic

control of vegetative bud set (Bohlenius *et al.* 2006). Similarly, a homologue of *FT* is proposed to play a role in the control of growth rhythm in Norway Spruce; *FT* upregulation is correlated with the induction of growth cessation and bud set under short days in autumn, and a decline in *FT* expression was associated with bud burst as temperatures increased in spring (Gyllenstrand *et al.* 2007). This alternative role of *FT* in a non-flowering plant could indicate that the ancestral function of *FT*-like genes was in the control of growth.

3.1.2.3 Flowering genes isolated in *Eucalyptus*

Homologues of the *Arabidopsis* flowering genes have been identified in *Eucalyptus*. Two homologues of *AP1* (*EAP1* and *EAP2*, though the latter is not a homologue of *Arabidopsis AP2* but rather a second homologue of *AP1*) and a homologue of *LEAFY* (*ELF1*) were identified in *Eucalyptus*, and transgenic expression of these genes in *Arabidopsis* showed that they are functionally equivalent to their *Arabidopsis* homologues (Kyozuka *et al.* 1997, Southerton *et al.* 1998b). *EgrSVP*, the *E. grandis* homologue of *SVP*, has been identified; its ectopic expression in *Arabidopsis* delayed the transition to flowering and altered the flowering phenotype in a variety of ways, including the homeotic conversion of the perianth into leaves, indeterminate flower development, and altered plant architecture (Brill and Watson 2004). Southerton *et al.* (1998a) identified three MADS-box genes in *E. grandis* (*EgM1*, *EgM2* and *EgM3*) that had strong homology to floral organ identity genes. *EgM2* was part of the *PISTILLATA* clade, genes of which function as "class B" genes in the ABC model of flower development. *EgM1* and *EgM3* had strongest homology to the *AGL2* (*SEPALLATA*) class E genes. The expression of these genes in sepals, petals, stamens and carpels of developing flowers was congruent with the gene homology, despite the modified floral structure of eucalypts (see 2.1.1, Figure 2.3). Dornelas and Rodriguez (2005a) identified many *Eucalyptus* homologues of *Arabidopsis* flowering time genes, including those mentioned above, but also, notably, including *FLC*. Unfortunately, these ESTs (from the Brazilian FORESTS database) are not yet publicly available.

While it has been shown that the *Eucalyptus* homologues of the *Arabidopsis* flowering pathway genes are expressed in specific tissue types, to date there have been no studies

of how the expression of these genes may be regulated during the processes of flower bud initiation, flower development and flower opening.

3.1.3 Aims of this study

This study focuses on the annual transition from vegetative to reproductive growth, flower bud development and flower opening in *globulus*, the last of which is under strong genetic control (see Chapter 2) and is economically significant. To identify the genes associated with seasonal flower bud initiation and opening in *globulus*, the homologues of several genes in the *Arabidopsis* flowering pathway were isolated from *globulus* and cloned. Their expression patterns were monitored in leaves, apices and flower buds of "early" (late autumn) and "late" (summer) flower opening genotypes, over a two year period.

3.2 Materials and Methods

3.2.1 Plant material

Tissue samples were harvested every two weeks from two ramets (i.e. two individual clones, separate trees) each of an early (late autumn) and late (summer) flower opening genotype (4489 and 5296, from the Furneaux and Strzelecki Ranges races respectively) from the Southern Tree Breeding Association (STBA) breeding population of *globulus*. These genotypes differ in their flower opening time, but vary only slightly in the timing of when flower buds are first observed (Table 3.2), so the difference in the observed flower opening time is due to different rates of development of the flower buds (Figure 2.9). These genotypes had been grafted onto a rootstock and planted at the seedEnergy Cambridge Orchard between September 2001 and May 2002, and treated at various ages with paclobutrazol.

Table 3.2. Tissue types harvested for each ramet at each sampling time point over the entire experiment. For tissue type codes see Table 3.3. An asterisk indicates that the tissue type was not harvested for that sample. An underlined sample indicates the timing of macroscopic appearance of flower buds for that ramet, shading indicates the period of flower opening.

			Genotype-ramet / column-row position in seed orchard							
			4489-1 / 10-36	4489-2 / 13-28	4489-3 / 27-26	4489-4 / 27-27	5296-1 / 11-04	5296-2 / 21-13	5296-3 / 12-34	5296-4 / 21-02
Year	Week	Date	"Early 1"	"Early 2"	"Early 3"	"Early 4"	"Late 1"	"Late 2"	"Late 3"	"Late 4"
1	0	17/05/06	*BCDE	ABCDE			ABCDE	*BCDE		
1	2	31/05/06	ABCDE	ABCDE	*BCDE		ABCDE	**CDE		
1	4	14/06/06		*BCDEF	*BCDEF		*BCDEF	*BCDEF		
1	6	28/06/06		*BCDEF	*BCDEF		*BCDEF	*BCDEF		
1	8	12/07/06		*BCDEF	*BC*EF		*BC*EF	*BCDEF		
1	10	26/07/06		*BCDEF	*BCDEF		*BC*EF	*BCDEF		
1	12	9/08/06		*BCDEF	**CDEF		*BC*EF	**CDEF		
1	14	28/08/06		*BCDEF	**CDEF		*BCDEF	**CDEF		
1	16	11/09/06		*BCDEF	*BCDEF		*BCDEF	*BCDEF		
1	18	27/09/06		*BCDEF	*BCDEF		*BCDEF	*BCDEF		
1	20	11/10/06		**CDEF	*BCDEF		*BCDEF	**CDEF		
2	22	25/10/06		**CDEFGH	*BCDEF		**CDEF	**CDEF	Flower initiation	
2	24	8/11/06		*BCDEFGH	*BCDEF		*BCDEF	*BCDEFGH		Flower initiation
2	26	22/11/06		*BCDEFGH	*BCDEF		*BCDEFGH	*BCDEFGH		
2	28	6/12/06		*BCDEFGH	*BCDEF		**CDEF*H	*BCDEFGH		
2	30	19/12/06		*BCDEFGH	*BCDEF		*BCDEFGH	**C*EFGH		
2	32	4/01/07		**CDE*GH	*BCDE*		*BCDEFGH	**CDEFGH		
2	34	18/01/07		***DE*GH	*BD*E*	**CDE*GH	*BC*E*GH	**cdE**H	**CDE*GH	*B*DE*GH
2	36	31/01/07		***DE*GH	*BCDE*	***DE*GH			***DE*GH	**CDE*GH
2	38	14/02/07		**CDE*GH	*BC*E*	***DE*GH	**C*E**H	***DE**H	**CDE*GH	**C*E*GH
2	40	28/02/07		**C*E*GH	*BCDE*	***DE*GH			**CDE*GH	*BC*E*GH
2	42	14/03/07		*BCDE*GHI	**CDE*	**C*E*GHI	**CDE*GHI	****E*GHI	**CDE*GHI	*BCDE*GHI
2	44	28/03/07		***DE*GHI	**CDE*	**C*E*GHI			**CDE*GHI	**CDE*GHI
2	46	12/04/07		*BCDE*GHI	*BCDE	**CDE**HI	*BCDE*GHI	***DE*GHI	**CDE*GHI	**CDE*GHI
2	48	26/04/07		*BCDE*GHI	**CDE	***DE**HI			**CDE*GHI	**CDE*GHI

2	50	10/05/07	***DE*GHI	**CDE	***DE**H*	**C*E*GHI	***DE*GHI	***DE*GHI	**CDE*GHI
2	52	24/05/07	***DE*GH*	**CDE	***DE**HI			**CDE*GHI	**C*E*GHI
2	54	7/06/07	***DE*GHI	**CDE	*B**E**H*	**CDE*GHI	***DE*GHI	**C*E*GHI	***DE*GHI
2	56	21/06/07	**CDE**H*	**CDE	***DE*GHI			**CDE*GHI	**CDE*GHI
2	58	4/07/07	****E*GHI	**CDE	***DE*GHI	****E*GHI	***DE*GHI	**CDE*GH*	**C*E*GHI
2	60	19/07/07	***DE*GHI	**CDE				***DE*GHI	***DE*GHI
2	62	1/08/07	***DE*GHI	**C*E	**CDE****	**CDE*GHI	***DE*GHI	*B*DE*GHI	***DE*GHI
2	64	15/08/07	***DE*GHI	*BCDE***I	*BC*E***I			**CDE*GHI	***DE*GHI
2	66	29/08/07	***DE*GHI	**CDE		**C*E**HI	***DE*GHI	**C*E*GHI	****E*GHI
2	68	13/09/07	***DE*GHI	**CDE***I	***DE*GHI			***DE*GHI	****E*GHIJ
2	70	27/09/07	***DE*GHIJ	**CDE****J	**C*E****J	**CDE*GH	***DE**HIJ	****E*GHI	****E*GHIJ
3	72	11/10/07	***DE***IJKL	**CDE*****KL	****E***I*KL			***DE*GHI*KL	****E*GHIJKL
3	74	24/10/07	****E*G**JKL	***DE*****KL	****E****JKL	****E**HIJKL	****E**HIJKL	**CDE**HI*KL	****E*GHI*KL
3	76	7/11/07	****E*G*IJKL	**CDE*****KL	***DE*****KL	****E*G*****		***DE*GH**KL	****E*GHIJKL
3	78	21/11/07	****E*G**JKL	**CDE*****KL	**C*E*****KL	****E**HI*KL	****E*GHI*KL	***DE*GHI*KL	****E*GH*JKL
3	80	5/12/07	**C*E****JKL	**C*E*****KL	**CDE*****KL	[]	[]	***DE*GH**KL	****E*GH*JKL
3	82	19/12/07	**C*E****JKL	**CDE*****KL	****E*****KLM	****E**H*JKL	****E**H**KL	****E*GHI*KLM	****E*GHI*KLM
3	84	3/01/08	**C*E****JKLM	**C*E****JKLM	**C*E*****KLM			***DE*GH**KLM	****E*GHI*KLM
3	86	16/01/08	****E****JKL	**C*E****JKL	**C*E*****KLM	****E*****KLM	****E****JKL	***DE*G***KL	****E*G*I*KLM
3	88	31/01/08	****E****JKLM	**CDE*****KL	**C*E*****KL			***DE*G*I*KL	****E*G*I*KLM
3	90	13/02/08	****E****JKLM	**C*E*****KL	**C*E*****KL	****E****JKLM	***DE****JKL	****E***I*KLM	****E*G**JKLM

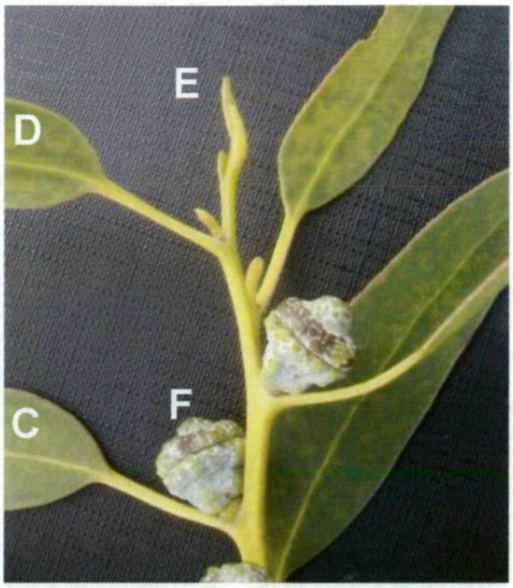
Shoot apices, inflorescence buds and leaves at different developmental stages were harvested (Table 3.3, Figure 3.2) and frozen in liquid nitrogen and stored at -80°C until RNA extraction. Samples of the same type and developmental stage were pooled into one sample for RNA extraction. For example, a single sample of leaf C in year one included 3-4 pooled leaves that had an inflorescence in the axil and originated from the same ramet and usually the same branch. Samples were harvested for 46 time points in total, covering two flowering seasons from winter 2006 to summer 2008 (Table 3.2). Tissues were sampled from the time of macroscopic appearance until the end of the study (week 90) where possible, and flower bud length and width were measured before harvesting.

Table 3.3. Eucalypt tissue types harvested for gene expression studies over two flowering seasons. The year the sample was first harvested is given, though samples were harvested until the end of the study where possible. For samples harvested in years two and three, the developmental equivalent in previous years is also given.

Sample	Year	Tissue	Tissue description	Equivalent in previous years
A	1	Leaf	With fruit in axil in year 1[1]	
B	1	Leaf	No inflorescence in axil; between A and C	
C	1	Leaf	With inflorescence F in axil in year 1 then fruit in year 2	
D	1	Leaf	No inflorescence in axil; between C and E	
E	1	Apex	Shoot tips, including unexpanded leaves	
F	1	Bud	In axil of leaf C	
G	2	Leaf	With inflorescence H in axil in year 2 then fruit in year 3	C
H	2	Bud	In axil of leaf G	F
I	2	Leaf	No inflorescence in axil; between G and E	D
J	3	Leaf	No inflorescence in axil; between I and E	
K	3	Leaf	With inflorescence L in axil in year 3	C, G
L	3	Bud	In axil of leaf K	F, H
M	3	Leaf	No inflorescence in axil; between K and E	J

[1] This tissue type was only sampled on the first harvest date.

a)



b)

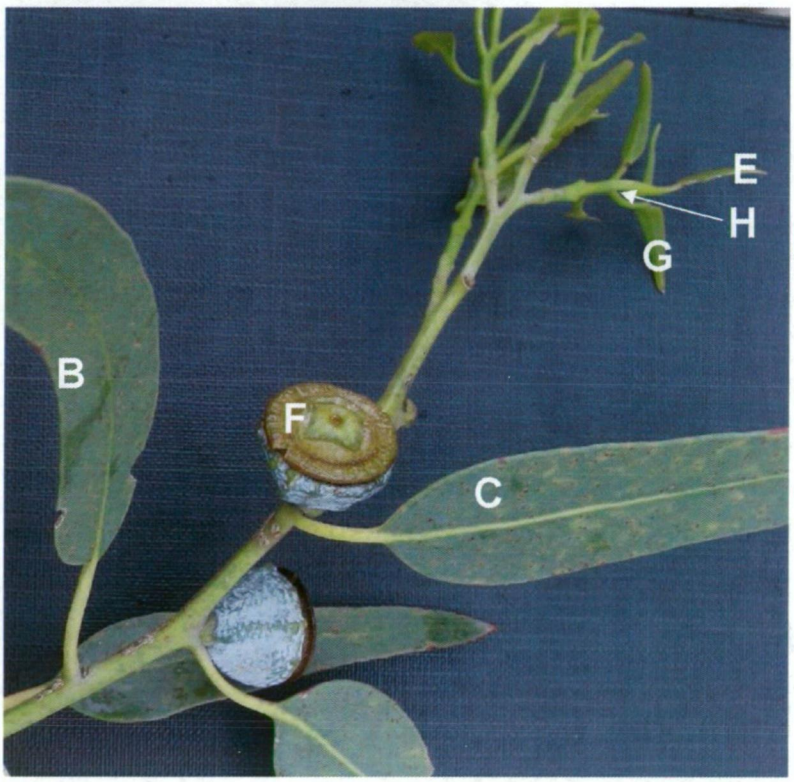


Figure 3.2. Eucalypt tissues harvested for gene expression studies over two flowering seasons. a) Typical year one (2006) sample, including flower buds (F) and leaves (C and D) initiated in spring 2005. b) Typical year two (2006-2007) sample, including developing fruit (F) and leaves (B and C) initiated spring 2005, and flower buds (H) and leaves (G) initiated in spring 2006.

Initially, two ramets of each genotype were harvested (4489-2, "Early 2"; 4489-3, "Early 3"; 5296-1, "Late 1"; 5296-2, "Late 2"), but as no flower buds were observed on Early 3 in spring 2006, an additional ramet (4489-4, "Early 4") was harvested. Due to a scarcity of available tissue for harvest in Late 1 and Late 2 by October 2006, two additional ramets (5296-3, "Late 3"; 5296-4, "Late 4") were harvested every two weeks and harvesting of Late 1 and Late 2 was reduced to every four weeks. All ramets were genotyped using the nine microsatellite markers used in Chapter 1 to verify that they were clonal.

Where possible, tissues were harvested at around 1 pm, but on occasion, some samples were taken later in the day. To test whether the abundance of key flowering and reference gene mRNA fluctuated during the potential sampling times, the leaves (G) and apices (E) of one ramet each of three different genotypes (4561, 5296, 5856) were harvested every two hours from 11 am to 5 pm, during vegetative flush and macroscopic appearance of flower buds in spring 2007.

3.2.2 Isolation of nucleic acids

Tissues were crushed in liquid nitrogen in a mortar and pestle prior to nucleic acid extraction. Nucleic acids were isolated using commercial kits with slight modifications to the standard protocols. Flower bud samples (F, H, L) and apex sample (E) RNA were isolated using the RNeasy mini kit (Qiagen), with 1% PVP (MW 40000) and 1% PEG (MW 20000) added to the RLT extraction buffer prior to use. Leaf (C, G, K) RNA was isolated using the SV Total RNA Isolation System (Promega), with 1% PEG (MW 20000) added to the lysis buffer. In both protocols, the DNase incubation was increased to 30 minutes. Genomic DNA (gDNA) was extracted using Plant DNA Isolation Kit (Mobio) with no modifications to the protocol. Nucleic acids were checked for purity on an agarose gel and quantified using the Picofluor Handheld Fluorometer (Turner BioSystems). The PicoGreen® and RiboGreen™ dyes were used for DNA and RNA quantification respectively, compared with a λDNA and 18S RNA standard respectively, according to manufacturer's instructions.

Double-stranded cDNA was synthesised using the Improm II RT kit (Promega), according to the protocol of the supplier, in a total volume of 20 µL. A negative control (i.e. omitting reverse transcriptase enzyme, RT-) was included for each RNA sample in order to check for genomic DNA contamination of RNA samples. The cDNA was diluted 5x with sterile distilled water and 5 µL was used in gene-specific PCRs for gene isolation and 2 µL was used in quantitative real-time PCRs (qRT-PCR).

3.2.3 Isolation of genes

To isolate the *Eucalyptus globulus* homologues of genes in the *Arabidopsis* flowering pathway and constitutively expressed reference genes, either published *Eucalyptus* sequences were used for primer design with Primer 3 v.0.4.0 (Rozen and Skaletsky 2000), or a number of plant sequences (Table 3.4) were aligned to design degenerate primers using Block Maker and CODEHOP (Rose *et al.* 1998).

Table 3.4. Sequences used to isolate *globulus* flowering and reference genes. Primer names and sequences given are those designed in this study. Expected fragment sizes are given in base pairs.

Gene	Homolog in *Arabidopsis*	Species	Reference	Accession	Primer name	Primer sequence	Expected size (cDNA)
EgFT	*FT*	*Arabidopsis thaliana*		NM_105222	FTdeg-1F	GMGAYGAYCTSCGGAMNTTYTAYAC	~350
		Populus tremula		DQ387859	FTdeg-1R	CCGCCRSAGCCGSWYTCNCKYTG	
		Glycine max		TC219541			
EgSOC1	*SOC1/*	*Arabidopsis thaliana*		NM_130128	SOC1deg-1F	GCCTTYGAGCTSTCCGTGYTNTGYGAYGC	~530
	AGL20	*Pisum sativum*		AY830920	SOC1deg-1R	KCCGATGAASAGMYCGGTNWCNACNTC	
		Citrus sinensis		EU032531			
		Glycine max		TC184543			
ELF1	*LFY*	*Eucalyptus globulus*	Southerton *et al.* 1998b	AF034806	EGELF-1F	GAGGAGCTGTTCGAGGCTTA	1181
					EGELF-1R	AAGGGAGTTCGAGATGGTGA	
EAP1	*AP1*	*Eucalyptus globulus*	Kyozuka *et al.* 1997	AF305076	EGAP1-1F	GTGTGAATTGGAGCGAGGAT	776
					EGAP1-1R	AGCAGCAGAAGTTCCCTTTT	
MeTFL1	*TFL1*	*Metrosideros excelsa*	Sreekantan *et al.* 2004	AY170872		See Sreekantan *et al.* 2004[1]	165
EgrSVP	*SVP*	*Eucalyptus grandis*	Brill and Watson 2004	AY263809	EGSVP-1F	TGGCGAGAGAAAAGATTCAGA	719
					EGSVP-1R	ATGCATCCCACCATTGTTGT	
EgM1	*SEP3*	*Eucalyptus grandis*	Southerton *et al.* 1998a	AF029975	EGM1-1F	CGAGTCCTTGGAAAGACAGC	105
					EGM1-1R	ACCTGTTCCCGATGTTGAAG	
EgM2	*PI*	*Eucalyptus grandis*	Southerton *et al.* 1998a	AF029976	EGM2-1F	GCTTTGGGATGCTAAGCAAG	185
					EGM2-1R	CCTTCTGGTCTCGGACACAT	
EgM3	*SEP1*	*Eucalyptus grandis*	Southerton *et al.* 1998a	AF029977	EGM3-1F	CAAGTGAACTGCTTGCTGGA	144
					EGM3-1R	GCCCTTCCTTTTAGGGTACG	
EgH4	*H4*	*Eucalyptus globulus*	Watson and Brill unpubl.	AY263810	EGH4-1F	GCGGCAAGGGAGGCAAGG	193
					EGH4-1R	CGGATCACGTTCTCCAGGAA	
EgTUBA1	*TUA6*	*Eucalyptus globulus*	Diaz *et al.* 1996	U37794	EGTB-1F	AGCGCCTGTCTGTGGATTAT	185
					EGTB-1R	AGAACGCCTGCAGATTTCAT	
18S	*18S*	*Pisum sativum*	Ozga *et al.* 2003			See Ozga *et al.* 2003	62
UBI	*UBI*	*Pisum sativum*	Albrecht *et al.* 1998			See Albrecht *et al.* 1998	3 bands
EF1	*EF1*	*Pisum sativum*	Foucher *et al.* 2003			See Foucher *et al.* 2003	213
ACT		*Pisum sativum*	Foo *et al.* 2005			See Foo *et al.* 2005	87
ACT11	*ACT11*	*Populus*	Brunner *et al.* 2004	CA824001		See Brunner *et al.* 2004	

Each PCR reaction mixture (25 µL final volume) contained 1x *Taq* polymerase reaction buffer (Fisher Biotech; 67 mM Tris-HCl pH 8.8, 16.6 mM $(NH_4)_2SO_4$, 0.45% Triton X-100, 0.2 mg/mL gelatin); 100 µg/mL of BSA (Bovine Serum Albumin); 2 mM $MgCl_2$; 160 µM each of dATP, dCTP, dGTP and dTTP; 7.5pmol of each primer; 5µL cDNA or gDNA and 2.2 units of *Taq* DNA polymerase, or, for some genes (see Table 3.4), Advantage ® Polymerase (Clontech). PCR amplification was performed in a PTC-100 Programmable Thermal Cycler or Tetrad Thermal Cycler (MJ Research, Inc.), using the following conditions: 94°C for 5 min; 30 cycles of 94°C for 1 min, optimum annealing temperature (see Table 3.4) for 1 min, 72°C for 2 min; and a final extension at 72°C for 10 min. For the genes that required Advantage ® Polymerase, the following conditions were used: 95°C for 1 min; 30 cycles of 95°C for 30 s, optimum annealing temperature (see Table 3.4) for 1 min, 68°C for 3 min; and a final extension at 68°C for 3 min.

PCR products were cleaned using the Wizard ® SV Gel and PCR Clean-Up System (Promega), quantified using a the Picofluor Handheld Fluorometer (Turner BioSystems) and PicoGreen® dye, and ligated into the pGEM®-T Easy Vector (Promega) according to manufacturer's instructions. Gene-specific PCR or pGEM primers (see below) were used to identify positive inserts and these were sequenced at the Australian Genome Research Facility (Brisbane, Australia) using ABI Prism™ BigDye™ Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems) with plasmid DNA as the template and primers specific to the pGEM®-T Easy Vector: pGEMT-F GCCCGACGTCGCATGCTCC and pGEMT-R GAGCTCTCCCATATGGTCG. Phylogenetic analyses were undertaken to confirm the identity of the genes that were isolated. The amino acid sequences were aligned with ClustalX (Larkin *et al.* 2007). A neighbour-joining tree was constructed in PAUP Version 4.0b10 (Swofford 2003) with 1000 bootstrap replicates.

3.2.4 Quantitative real-time PCR

3.2.4.1 Quantitative real-time PCR reaction conditions

Primers for qRT-PCR were designed based on the sequences obtained in 3.2.3, binding to separate exons when both gDNA and cDNA sequences were available, to avoid

generating false positives from gDNA contamination (Bustin 2000). Where the sequenced PCR product length was less than 200 bp, the same primers were used (Table 3.5). Reactions contained 1x SensiMix (Quantace), 1x SYBR ® green solution, 0.3 μM each primer and 2 μL each cDNA sample (~ 10 ng) in a total volume of 10 μL. Real-time PCR was performed on the Rotor-Gene™ 3000 (Corbett) using the following conditions: 95°C for 10 min; 50 cycles of 95°C for 5 s and 60°C for 40 s; 65°C for 45 s then a ramp from 65°C to 95°C in one degree increments, with 5 s at each step. Each reaction was performed in duplicate and no-template controls were included. In addition, seven samples of plasmid DNA with an insert of the gene of interest, in a 10-fold dilution series from 1 ng/μL to $1x10^{-6}$ ng/μL were included. The constitutive genes *EgH4*, *EgTUBA1*, *Eg18S*, *EgUBI*, *EgEF1* and *EgACT11* were tested as reference genes in leaf and apex tissue of each genotype, and *Eg18S* was tested in bud tissue of Early 2.

Table 3.5. Primers for qRT-PCR of *globulus* flowering genes.

Gene	Homolog	Primer names	Primer sequences
EgFT	*FT*	EGFT-2F	TGATATTCCAGCTACGACAGGA
		EGFT-2R	CCCCAGGTTGTAGAGCTCAG
EgSOC1	*SOC1/*	EGSOC1-2F	GCTGCGAACATGATGAAGAA
	AGL20	EGSOC1-2R	ATGCTGATAACGCTCCGTTC
ELF1	*LFY*	EGELF-2F	AACGGCCTGGACTACCTCTT
		EGELF-2R	GCAGTGGACATAGTGCCTCA
EAP1	*AP1*	EGAP1-2F	TGCTAAGCTCAAGGCCAGAC
		EGAP1-2R	GCAGACTCCAGTTGCTGCTC
EgTFL1	*TFL1*	EGTFL1-5F	CAGACCCAGATGTTCCTGGT
		EGTFL1-4R	CGAACCTGTGGATACCAATG[1]
EgSVP	*SVP*	EGSVP-2F	TTGCAGAGAAAGGTCACCAA
		EGSVP-2R	TGTTGAAGGTCGGTGATCTCT
EgM1	*SEP3*	EGM1-1F	As in Table 3.4
		EGM1-1R	
EgM2	*PI*	EGM2-1F	As in Table 3.4
		EGM2-1R	
EgM3	*SEP1*	EGM3-1F	As in Table 3.4
		EGM3-1R	
EgH4	*H4*	EGH4-1F	As in Table 3.4
		EGH4-1R	
EgTUBA1	*TUA6*	EGTB-1F	As in Table 3.4
		EGTB-1R	
Eg18S	*18S*		See Ozga *et al.* 2003
EgUBI	*UBI*		See Albrecht *et al.* 1998
EgEF1	*EF1*		See Johnson *et al.* 2006
EgACT11	*ACT11*		See Brunner *et al.* 2004[1]

[1]This primer is one of the nested primers from Sreekantan *et al.* 2004

3.2.4.2 Quantitative real-time PCR data analysis

Results were excluded if one of the duplicate reactions failed, and if the equivalent reaction using the template produced with reverse transcriptase omitted (RT-) had a threshold cycle (C_T, Pfaffl 2001) of less than 23 (i.e. gDNA contamination) or for a *Eg18S* $C_T > 20.0$ (i.e. low cDNA template). Duplicate C_Ts were averaged and relative transcript levels were evaluated using the mathematical model in Pfaffl (2001), using *Eg18S* as the reference gene.

3.3 Results

3.3.1 Gene isolation and qRT-PCR primer design

The flowering genes *EgFT*, *EgTFL1*, *EgSOC1*, *ELF1*, *EAP1*, *EgSVP*, *EgM1*, *EgM2* and *EgM3* were all isolated successfully from *globulus* cDNA (Table 3.6). The candidate reference genes *EgH4*, *EgTUBA1*, *Eg18S*, *EgUBI*, *EgEF1* and *EgACT11* were also isolated (Table 3.6). All PCR fragments were cloned in the pGEM®-T Easy Vector and sequenced. These sequences were used to design primers for qRT-PCR and in some cases the primers used for gene isolation could also be used for qRT-PCR (Table 3.5, Table 3.7).

Table 3.6. Optimal PCR conditions and sequence lengths (in base pairs) for the *globulus* genes isolated in this study. Expected and observed fragment sizes are given in base pairs. T_a, annealing temperature.; n.a., not amplified.

Gene family	Gene	Expected size (cDNA)	T_a	Observed size	
				gDNA	cDNA
FT/TFL1	*EgFT*	~350	55	1007	343
	EgTFL1	165	48, 51[3]	627	161
MADS	*EgSOC1*	~530	55	>1300[1]	518
	EAP1	776	65[2]	n.a.	776
	EgSVP	719	55	n.a.	718
	EgM1	105	60	n.a.	~100
	EgM2	185	60	n.a.	~180
	EgM3	144	60	n.a.	~150
LFY	*ELF1*	1181	55[1]	1182	1023
Constitutive	*EgH4*	193	60	193	193
	EgTUBA1	185	60	185	185
	Eg18S	62	60	~80	~80
	EgUBI	3 bands	60	multi-banded	multi-banded
	EgEF1	213	60	~300	~200
	EgACT	87	n.a.	n.a.	n.a.
	EgACT11		50	~200	~200

[1] Incomplete gDNA sequence was obtained. [2] Advantage ® Polymerase used. [3] PCR with annealing temperature of 48°C followed by nested PCR with annealing temperature of 51°C as in Sreekantan *et al.* 2004.

Table 3.7. Expected fragment sizes (in base pairs) and optimal qRT-PCR conditions for the *globulus* genes isolated in this study. T_a, annealing temperature.; n.d. not determined.

Gene family	Gene	Size		T_m range	Mean reaction efficiency
		gDNA	cDNA		
FT/TFL1	*EgFT*	331	190	86.5-88.3	1.73
	EgTFL1	619	153	84.0-85.7	1.74
MADS	*EgSOC1*	n.d.	134	81.0-82.7	1.76
	EAP1	n.d.	120	79.2-81.0	1.77
	EgSVP	n.d.	163	81.5-82.3	1.77
	EgM1	n.d.	~100	79.8-80.8	1.75
	EgM2	n.d.	~180	81.3-82.7	1.73
	EgM3	n.d.	~150	79.0-80.3	1.75
LFY	*ELF1*	282	187	86.0-87.3	1.74
Constitutive	*EgH4*	193	193	91.2-91.8	1.76
	EgTUBA1	185	185	84.7-85.3	1.75
	Eg18S	~80	~80	81.5-87.7	1.70
	EgUBI	multi-banded	multi-banded	85.8-88.2	1.72
	EgEF1	~300	~200	83.7-85.0	1.76

3.3.1.1 *FT/TFL1* gene family

3.3.1.1.1 FLOWERING LOCUS T (FT)

EgFT was isolated from gDNA and leaf cDNA using degenerate primers designed using *Arabidopsis thaliana*, *Populus tremula* and *Glycine max FT* sequences (Table 3.4, Table 3.6). Only one *FT* was isolated in *globulus* using this method, while studies in other forest tree species have isolated more than one *FT* homologue per species. *EgFT* clustered within the *FT* clade of the *FT/TFL1* gene family (Figure 3.3). Additional non-degenerate primers were designed for qRT-PCR of *EgFT* (Table 3.5) and were successful at an annealing temperature of 60°C with a high reaction efficiency and a single product (i.e. a narrow T_m range) (Table 3.7).

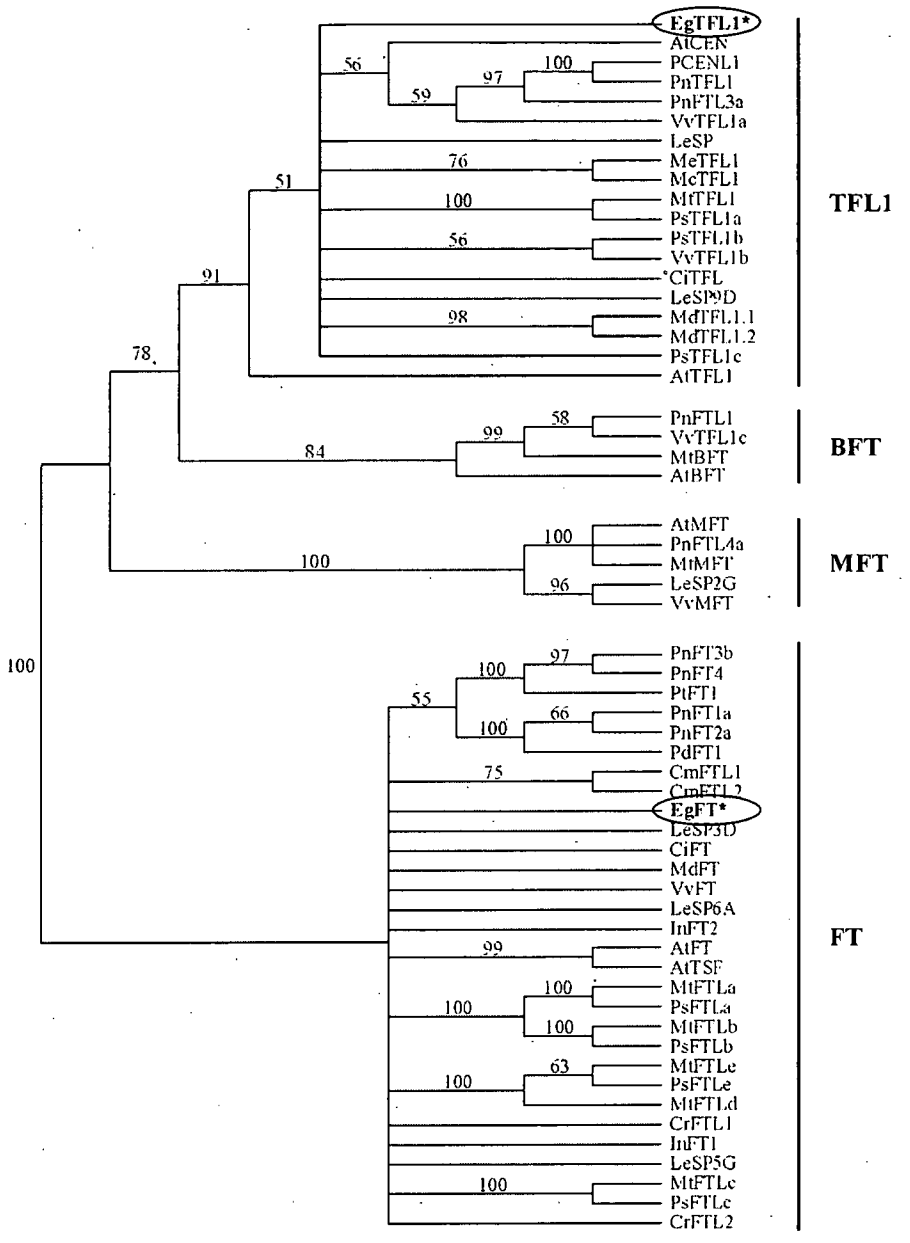


Figure 3.3. Phylogenetic analysis of several plant *FT/TFL1*-like genes and the *globulus* putative homologues of *FT* and *TFL1* (*EgFT* and *EgTFL1*). Neighbour-joining tree for the amino acid sequence aligned with ClustalX. Bootstrap values are indicated as a percentage above each branch. Genes isolated in this study and for which the expression patterns are studied are circled and marked by an asterisk.

At, *Arabidopsis thaliana*; Ci, *Citrus unshiu*; Cm, *Cucurbita maxima*; Cr, *Chenopodium rubrum*; Eg, *Eucalyptus globulus*; In, *Ipomoea nil*; Le, *Lycopersicon esculentum*; Mc, *Metrosideros collina*; Me, *M. excelsa*; Md, *Malus* x domestica; Mt, *Medicago truncatula*; Pd, *Populus deltoides*; Pn, *P. nigra*; Pt, *P. tremula*; P, *P. trichocarpa*; Ps, *Pisum sativum*; Vv, *Vitis vinifera*.

3.3.1.1.2 TERMINAL FLOWER1 (TFL1)

EgTFL1 was isolated from gDNA and apex cDNA using degenerate primers and nested PCR, as detailed by Sreekantan *et al.* (2004) (Table 3.4, Table 3.6). *EgTFL1* had close affinities to other plant *TFL1* genes (Figure 3.3). An additional non-degenerate forward primer was designed for use in combination with the reverse nested primer from Sreekantan *et al.* (2004) for qRT-PCR of *EgTFL1* (Table 3.4). This combination was successful at an annealing temperature of 60°C with a high reaction efficiency and a single product (i.e. a narrow T_m range) (Table 3.6).

3.3.1.2 MADS-box gene family

3.3.1.2.1 SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1 (SOC1)

The two *SOC1*-like sequences published for *E. grandis* (*EgrMADS3* and *EgrMADS4*, Watson and Brill 2004) had closest homology to *Arabidopsis AGL42* rather than *AGL20* (*AtSOC1*), and so degenerate primers were used to isolate the *globulus* homologue of *AGL20* (*SOC1*) (Table 3.4). *EgSOC1* was isolated from gDNA and leaf cDNA using degenerate primers designed using *Arabidopsis thaliana*, *Pisum sativum*, *Citrus sinensis* and *Glycine max SOC1* sequences (Table 3.4, Table 3.6). The *EgSOC1* cDNA sequence obtained in this study had closer affinities to *AtSOC1* than the other *Eucalyptus SOC1*-like proteins that have been isolated previously (Watson and Brill 2004) (Figure 3.4). Additional non-degenerate primers were designed for qRT-PCR of *EgSOC1* (Table 3.5) and were successful at an annealing temperature of 60°C with a high reaction efficiency and a single product (i.e. a narrow T_m range) (Table 3.7).

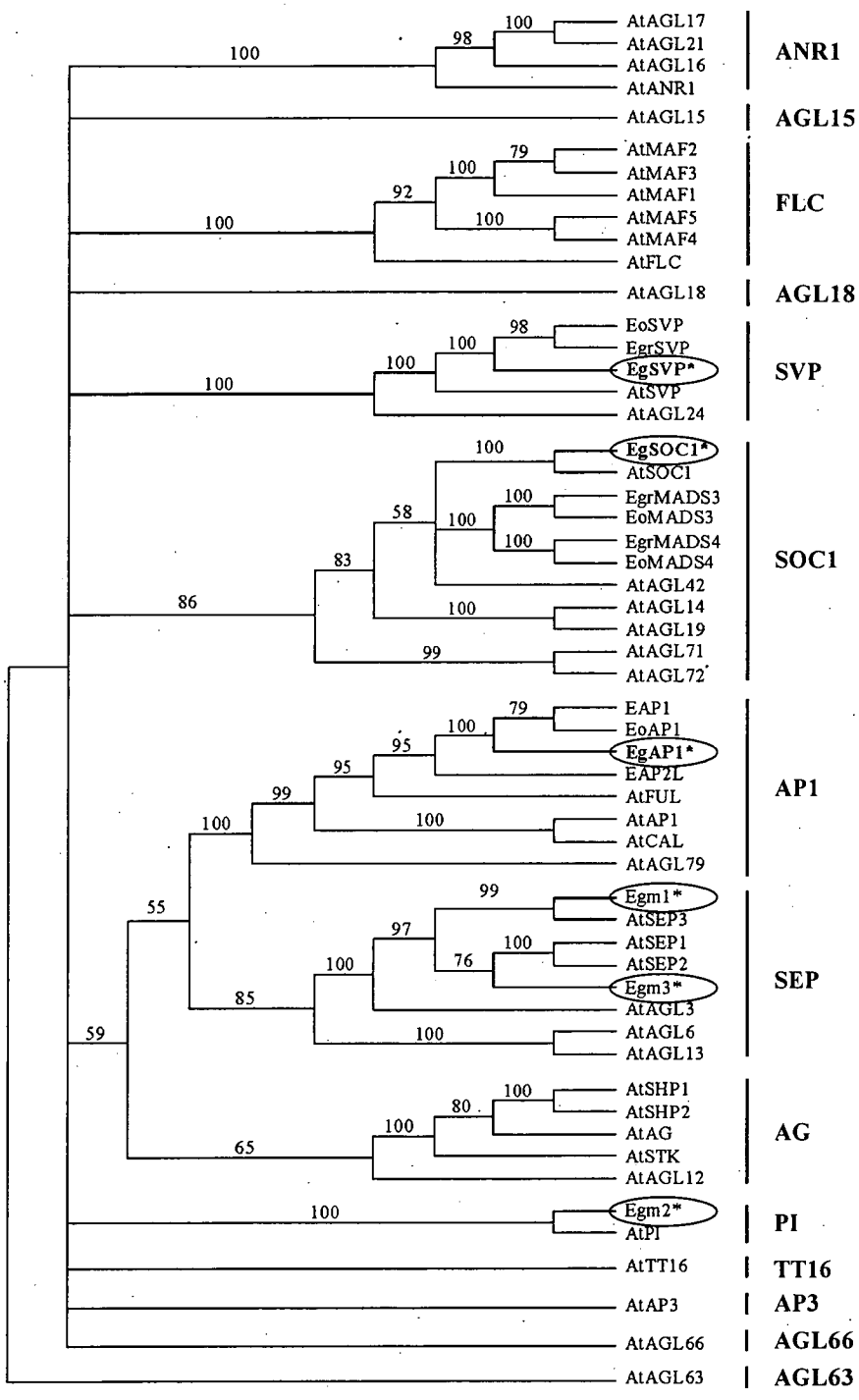


Figure 3.4. Phylogenetic analysis of *Arabidopsis* and *Eucalyptus* MIKC MADS-box genes. Neighbour-joining tree for the amino acid sequence aligned with ClustalX. Bootstrap values are indicated as a percentage above each branch. Genes isolated in this study are highlighted in bold and genes for which the expression patterns are studied here are circled and marked by an asterisk. At, *Arabidopsis thaliana*; E, *Eucalyptus*.

3.3.1.2.2 APETALA1 (AP1)

EgAP1 was isolated from *globulus* apex cDNA (Table 3.6) with primers designed using the *EAP1* sequence from *Eucalyptus globulus* ssp. *bicostata* (Kyozuka *et al.* 1997) (Table 3.4), but the gene could not be amplified from gDNA, despite using Advantage ® Polymerase and up to a 5 minute extension time in PCR cycles (Table 3.6). Attempts to amplify the gDNA of *EgAP1* in two parts using a combination of the original and qRT-PCR primers (i.e. EGAP1-1F/EGAP1-2R and EGAP1-2F/EGAP1-1R) were also unsuccessful (Table 3.6). *EgAP1* fell within the AP1 clade of MIKC MADS-box genes (Figure 3.4). Additional primers were designed for qRT-PCR of *EgAP1* (Table 3.5) and were successful at an annealing temperature of 60°C with a high reaction efficiency and a single product (i.e. a narrow T_m range) (Table 3.7).

3.3.1.2.3 SHORT VEGETATIVE PHASE (SVP)

EgSVP was isolated from *globulus* apex cDNA (Table 3.6) with primers designed using the *EgrSVP* sequence from *Eucalyptus grandis* (Brill and Watson 2004) (Table 3.4), but the gene could not be amplified from gDNA, despite using the highly sensitive, high fidelity enzyme Advantage ® Polymerase and up to a 5 minute extension time in PCR cycles (Table 3.6). Attempts to amplify the gDNA of *EgSVP* in two parts using a combination of the original and qRT-PCR primers (i.e. EGSVP-1F/EGSVP-2R and EGSVP-2F/EGSVP-1R) were also unsuccessful (Table 3.6). *EgSVP* fell within the SVP clade of MIKC MADS-box genes (Figure 3.4). Additional primers were designed for qRT-PCR of *EgSVP* (Table 3.5) and were successful at an annealing temperature of 60°C with a high reaction efficiency and a single product (i.e. a narrow T_m range) (Table 3.7).

3.3.1.2.4 SEPALLATA (SEP)

EgM1 was isolated from *globulus* flower bud cDNA (Table 3.6) with primers designed using the sequence of *EgM1* from *Eucalyptus grandis* (Southerton *et al.* 1998a) (Table 3.4), but the gene was not sequenced. The original sequence from Southerton *et al.* (1998a) fell within the SEP clade of MIKC MADS-box genes (Figure 3.4). The same primers were used for qRT-PCR of *EgM1* (Table 3.5) and were successful at an

annealing temperature of 60°C with a high reaction efficiency and a single product (i.e. a narrow T_m range) (Table 3.7).

EgM3 was isolated from *globulus* flower bud cDNA (Table 3.6) with primers designed using the sequence of *EgM3* from *Eucalyptus grandis* (Southerton *et al.* 1998a) (Table 3.4), but the gene was not sequenced. The original sequence from Southerton *et al.* (1998a) fell within the SEP clade of MIKC MADS-box genes (Figure 3.4). The same primers were used for qRT-PCR of *EgM3* (Table 3.5) and were successful at an annealing temperature of 60°C with a high reaction efficiency and a single product (i.e. a narrow T_m range) (Table 3.7).

3.3.1.2.5 PISTILLATA (PI)

EgM2 was isolated from *globulus* flower bud cDNA (Table 3.6) with primers designed using the sequence of *EgM2* from *Eucalyptus grandis* (Southerton *et al.* 1998a) (Table 3.4), but the gene was not sequenced. The original sequence from Southerton *et al.* (1998a) fell within the PI clade of MIKC MADS-box genes (Figure 3.4). The same primers were used for qRT-PCR of *EgM2* (Table 3.5) and were successful at an annealing temperature of 60°C with a high reaction efficiency and a single product (i.e. a narrow T_m range) (Table 3.7).

3.3.1.3 *LEAFY (LFY)*

EgLFY was isolated from *globulus* gDNA and apex cDNA with primers designed using the *globulus ELF1* sequence (Southerton *et al.* 1998b) (Table 3.4, Table 3.6). The sequence obtained had strong homology with *AtLFY* and *ELF1* (data not shown). Additional primers were designed for qRT-PCR of *EgLFY* (Table 3.5) and were successful at an annealing temperature of 60°C with a high reaction efficiency and a single product (i.e. a narrow T_m range) (Table 3.7).

3.3.1.4 Constitutive genes

3.3.1.4.1 Histone (*H4*)

EgH4 was isolated from *globulus* gDNA and flower bud cDNA with primers designed using the *globulus EgH4* sequence (Watson and Brill unpubl., Table 3.4, Table 3.6).

The same primers were used for qRT-PCR of *EgH4* (Table 3.5) and were successful at an annealing temperature of 60°C with a high reaction efficiency and a single product (i.e. a narrow T_m range) (Table 3.7).

3.3.1.4.2 RNA subunit *18S*

Eg18S was isolated from *globulus* gDNA and apex cDNA with primers used in *Pisum sativum* (Ozga *et al.* 2003, Table 3.4, Table 3.6). The same primers were used for qRT-PCR of *Eg18S* (Table 3.5) and were successful at an annealing temperature of 60°C with a high reaction efficiency but with a wide T_m range (Table 3.7).

3.3.1.4.3 α-TUBULIN (TUA)

EgTUBA1 was isolated from *globulus* gDNA and apex cDNA with primers designed using the *globulus EgTUBA1* sequence (Diaz *et al.* 1996, Table 3.4, Table 3.6). The same primers were used for qRT-PCR of *EgTUBA1* (Table 3.5) and were successful at an annealing temperature of 60°C with a high reaction efficiency and a single product (i.e. a narrow T_m range) (Table 3.7).

3.3.1.4.4 UBIQUITIN (UBI)

EgUBI was isolated from *globulus* gDNA and apex cDNA with primers used in *Pisum sativum* (Albrecht *et al.* 1998, Table 3.4, Table 3.6). The same primers were used for qRT-PCR of *EgUBI* (Table 3.5) and were successful at an annealing temperature of 60°C with a high reaction efficiency but with a wide T_m range (Table 3.7). Despite producing a multi-banded product, the *UBI* primers are used widely to amplify this reference gene in studies of gene expression in pea, as all bands are amplified quantitatively (J. Smith UTAS pers. comm.).

3.3.1.4.5 ELONGATION FACTOR 1 (EF1)

EgEF1 was isolated from *globulus* gDNA and apex cDNA with primers used in *Pisum sativum* (Foucher *et al.* 2003, Table 3.4, Table 3.6). The same primers were used for qRT-PCR of *EgEF1* (Table 3.5) and were successful at an annealing temperature of 60°C with a high reaction efficiency and a single product (i.e. a narrow T_m range) (Table 3.7).

3.3.1.4.6 *ACTIN (ACT)*

The *Pisum sativum* actin primers (Foo *et al.* 2005) did not amplify the *globulus* actin homolog from either cDNA or gDNA, but the *Populus* primers (Brunner *et al.* 2004) were successful (Table 3.6). The *Populus* primers were used for qRT-PCR of *EgACT11* (Table 3.5) but these primers did not amplify a product.

3.3.2 *Expression patterns of reference genes*

To determine which reference gene was most appropriate, the abundance of *EgH4*, *EgTUBA1*, *Eg18S*, *EgUBI*, *EgEF1* transcripts across nearly a year of harvests was measured and a regression line fitted. In leaf tissue, *Eg18S* was the most abundant (i.e. lowest C_T) reference gene and *EgTUBA1* was the least abundant (Figure 3.5). All reference genes showed a more or less consistent expression among samples of leaf tissue, with a low slope and close fit to the regression line (Figure 3.5a, c). *Eg18S* was also expressed at high levels in apex tissue, but *EgUBI* and *EgTUBA1* were the least abundant transcript in this tissue type (Figure 3.5). All of the reference genes were less stably expressed across different samples of apex tissue than they were in leaves (Figure 3.5) but the pattern was consistent across genes for a given harvest date. The expression profile of *Eg18S* in bud tissue was also consistent (Figure 3.6) and as this gene also had a stable expression pattern in leaves and apices (Figure 3.5), it was chosen to normalise data in subsequent experiments.

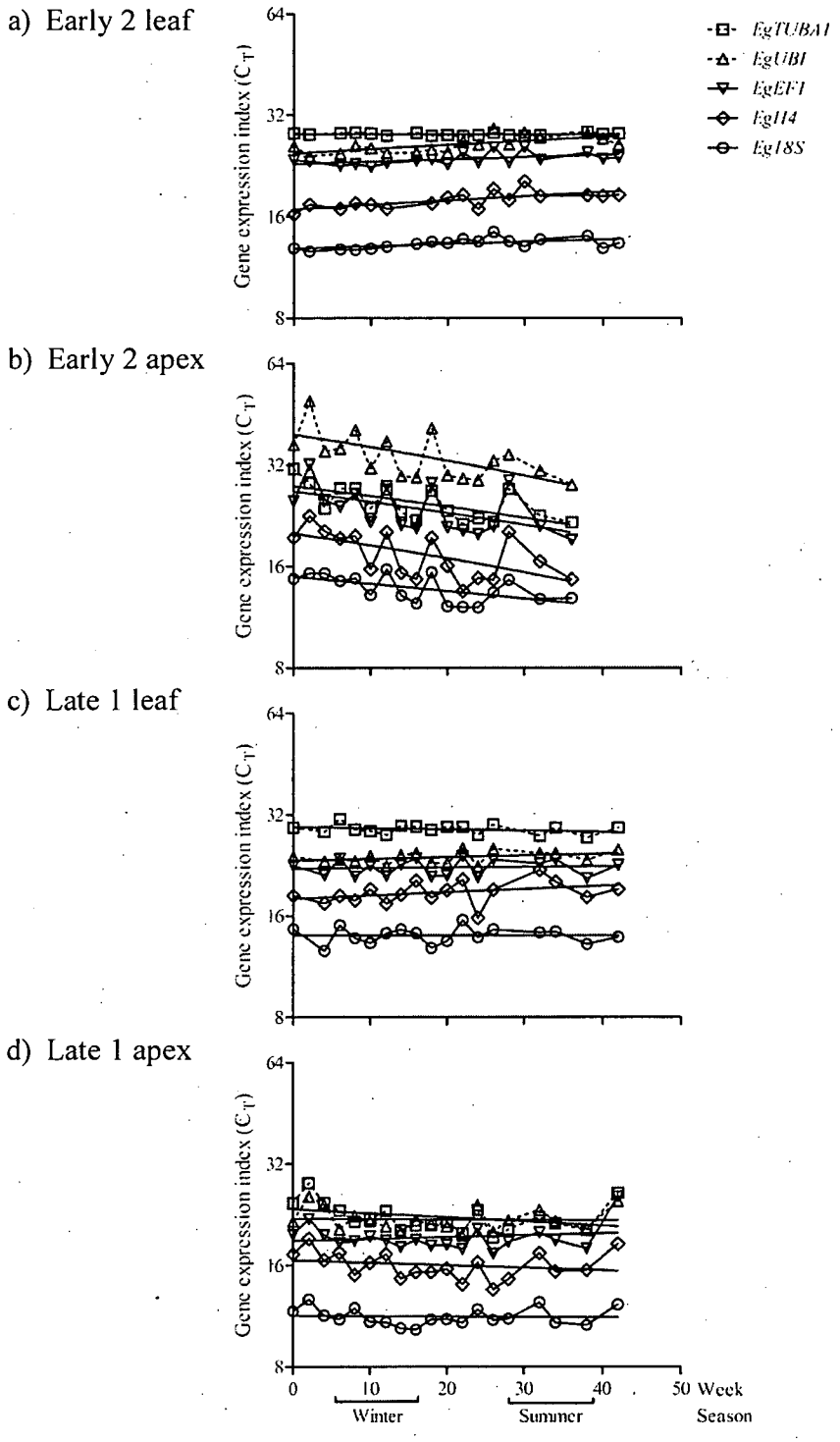


Figure 3.5. Variation in expression of reference genes over 46 weeks, showing regression lines for leaf and apex tissue in early and late genotypes of *globulus*: a) 4489-2 leaf C tissue; b) 4489-2 apex E tissue; c) 5296-1 leaf C tissue; d) 5296-1 apex E tissue.

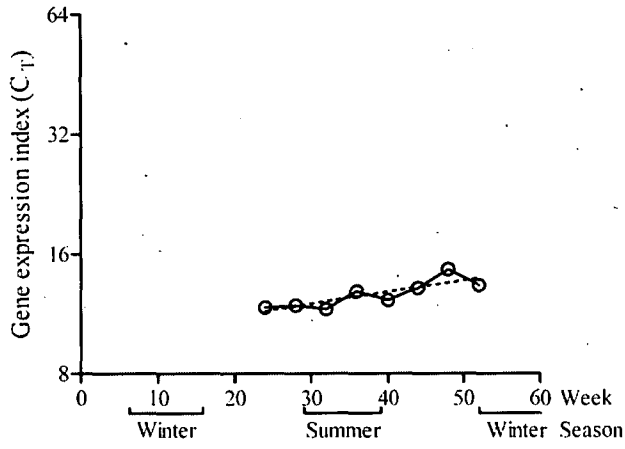
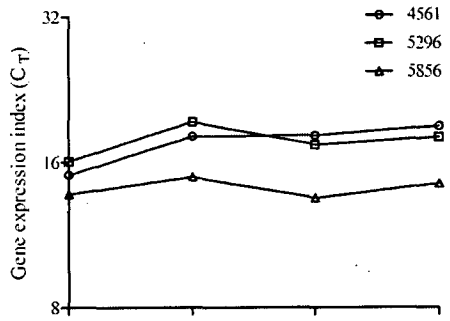


Figure 3.6. Variation in expression of the reference gene *Eg18S* over 14 weeks in buds of 4489-2, an early flower opening genotype of *globulus*. Regression line is shown by a dashed line.

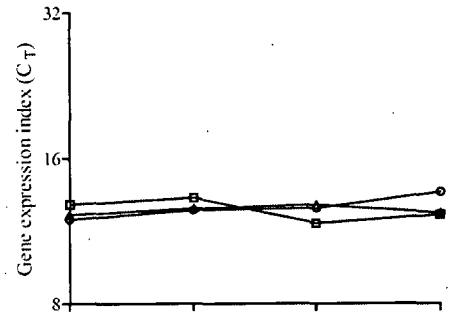
3.3.3 Diurnal expression patterns

To test whether the abundance of key flowering and reference gene mRNA fluctuated throughout the potential sampling times, *Eg18S*, *EgFT* and *ELF1* transcript levels were measured every two hours from 11 am to 5 pm. *EgFT* was measured only in leaf tissue as it was originally isolated from leaf but not apex tissue (see 3.3.1.1) and *ELF1* was measured only in apex tissue as it was originally isolated from apex but not leaf tissue (see 3.3.1.3). Expression of the reference gene *Eg18S* was relatively constant from 11 am to 5 pm in leaf and apex tissue (Figure 3.7a, b). The *EgFT* and *ELF1* transcript levels varied throughout this sampling period, but with no clear pattern that was consistent among genotypes (Figure 3.7c, d).

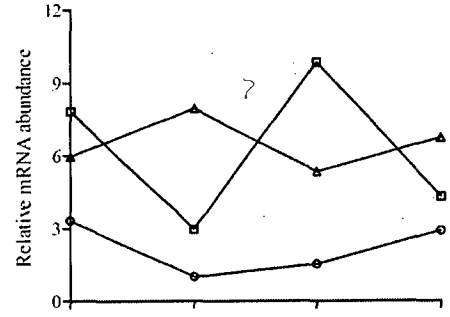
a) *Eg18S* leaf



b) *Eg18S* apex



c) *EgFT* leaf



d) *ELF1* apex

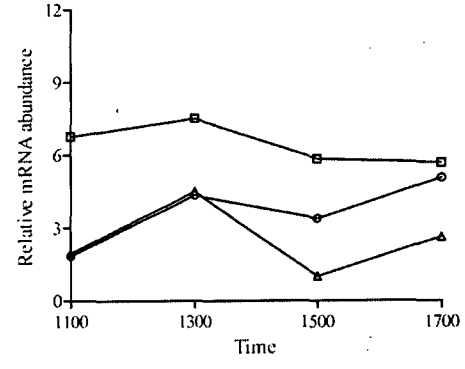


Figure 3.7. Variation in expression of reference and flowering genes in three genotypes of *globulus* harvested from 11am to 5pm during flower bud initiation in spring 2007: a) *Eg18S* in leaf G; b) *Eg18S* in apex E; c) *EgFT* mRNA abundance, relative to *Eg18S*, in leaf G; d) *ELF1* mRNA abundance, relative to *Eg18S*, in apex E.

3.3.4 Expression patterns of flowering genes

To test whether key *Arabidopsis* flowering gene homologues were associated with the seasonal initiation of flower buds and flower opening in *globulus*, the abundance of the flowering gene transcripts were compared in early (Early 2, Early 3, Early 4) and late (Late 1, Late 2) flower opening genotypes. The flowering genes studied represented pathway integrator (*FT*), repressor (*SVP*, *TFL1*), inflorescence identity (*LFY*, *SOC1*, *AP1*) and flower organ identity (*SEP*, *PI*) roles in the *Arabidopsis* flowering pathway.

3.3.4.1 Pathway integration genes

3.3.4.1.1 EgFT

Expression of *EgFT* in the young leaves of each year was low in late summer through autumn to early winter and then gradually increased during the transition to flower bud initiation in spring of each year. This pattern was most obvious in the late flower opening genotype (Late 1, Figure 3.8c; Late 2, Figure 3.8d) for which the timing of flower bud initiation and flower bud opening coincided.

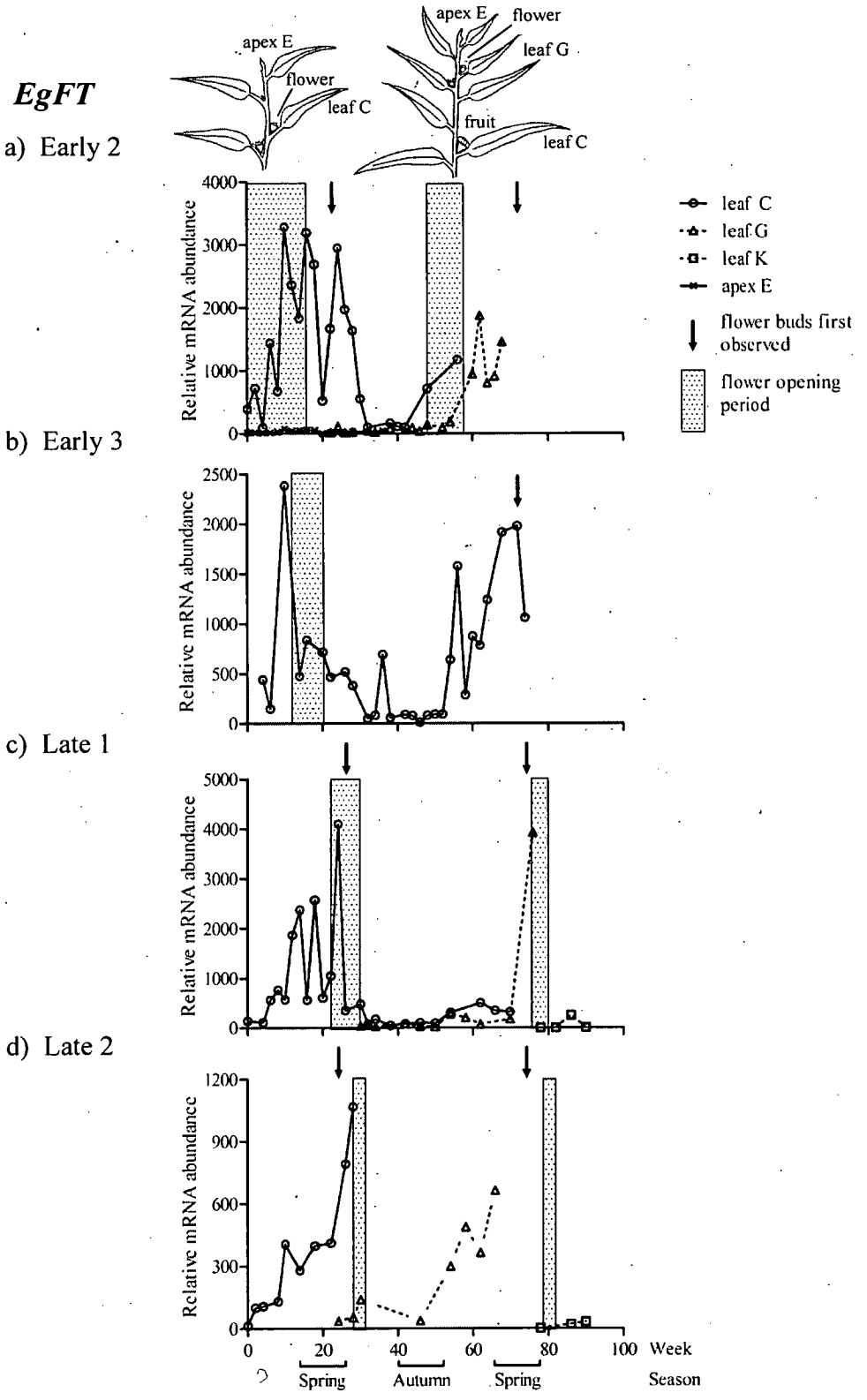


Figure 3.8. Relative *EgFT* mRNA abundance in leaves (C, G, K) and apices (E) of two ramets each of an early and late flower opening genotype of *globulus*: a) 4489-2; b) 4489-3; c) 5296-1; d) 5296-2. Timing of flower bud initiation and opening for each ramet is indicated, and representative shoot architecture of 4489-2 at the beginning of each year is also shown.

Comparison of the early and late flower opening clones showed no association between the expression of *EgFT* and the timing of flower bud opening, as the timing of *EgFT* expression in Early 2 was similar to that of the Late ramets (Figure 3.8a) despite a six month difference in the timing of flower bud opening between the two genotypes (Table 3.2). *EgFT* expression was therefore associated with flower bud initiation but not flower opening time.

Early 3 had a brief peak in *EgFT* expression in the first year leaves, but this was not ongoing as in the other ramets (Figure 3.8). There was, however, an ongoing upregulation of *EgFT* in these leaves in the second year (Figure 3.8b). There were no buds initiated in this ramet in the first year and therefore it did not have a flower opening period in the second year, but buds were initiated on this ramet in the second year (Table 3.2). *EgFT* expression was very low in the Early 2 apex samples across nearly a full year of harvests (Figure 3.8a), and was therefore not measured in the apex tissue of the other three ramets.

3.3.4.2 Repressor genes

3.3.4.2.1 EgSVP

In the first year, expression of *EgSVP* in Early 2 peaked in late summer, in both the apex (Figure 3.9a) and flower buds (Figure 3.10a,b). This coincided with the onset of winter dormancy in this tree. *EgSVP* expression also peaked at week 38 in Late 1 bud tissue (Figure 3.10c,d) but this pattern was not observed in the Late 1 apex (Figure 3.9b). Expression of *EgSVP* was very low in leaf tissue of Early 2 across nearly a full year of harvests (Figure 3.9a), so it was not measured in the leaf tissue of Late 1.

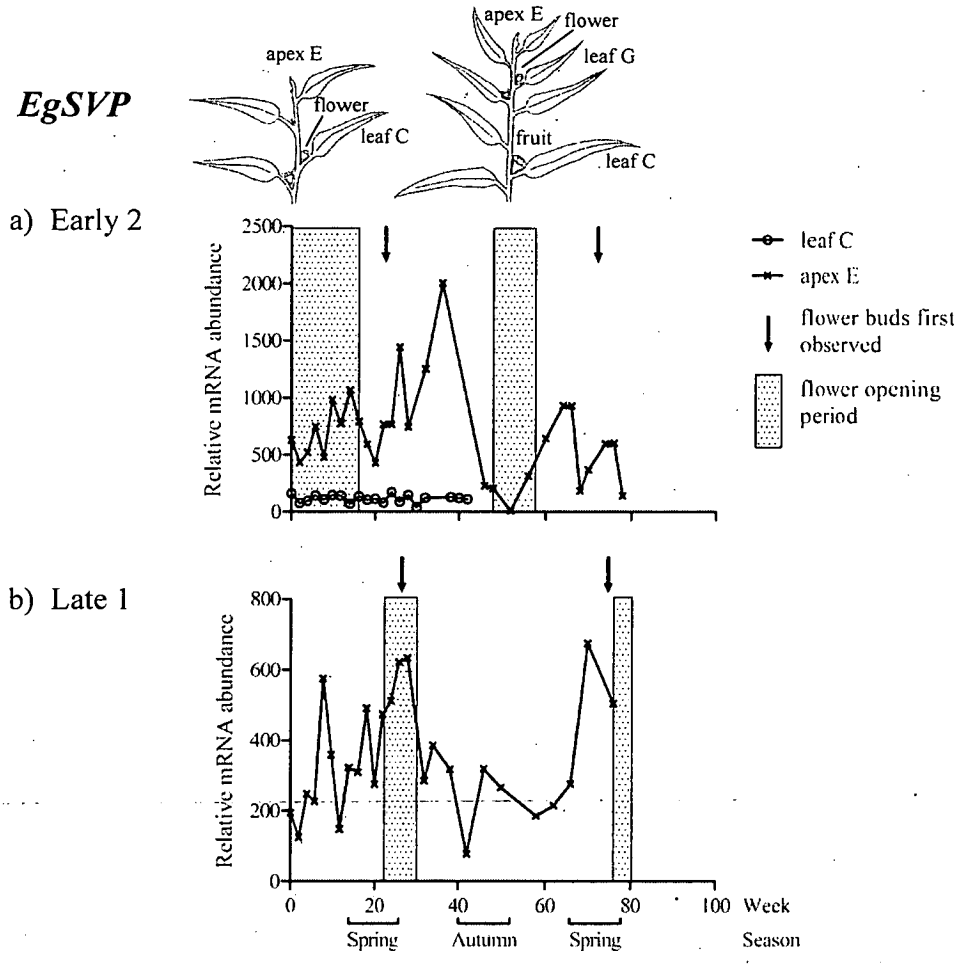


Figure 3.9. Relative *EgSVP* mRNA abundance in leaves (C) and apices (E) of one ramet each of an early and late flower opening genotype of *globulus*: a) 4489-2; b) 5296-1. Timing of flower bud initiation and opening for each ramet is indicated, and representative shoot architecture of 4489-2 at the beginning of each year is also shown.

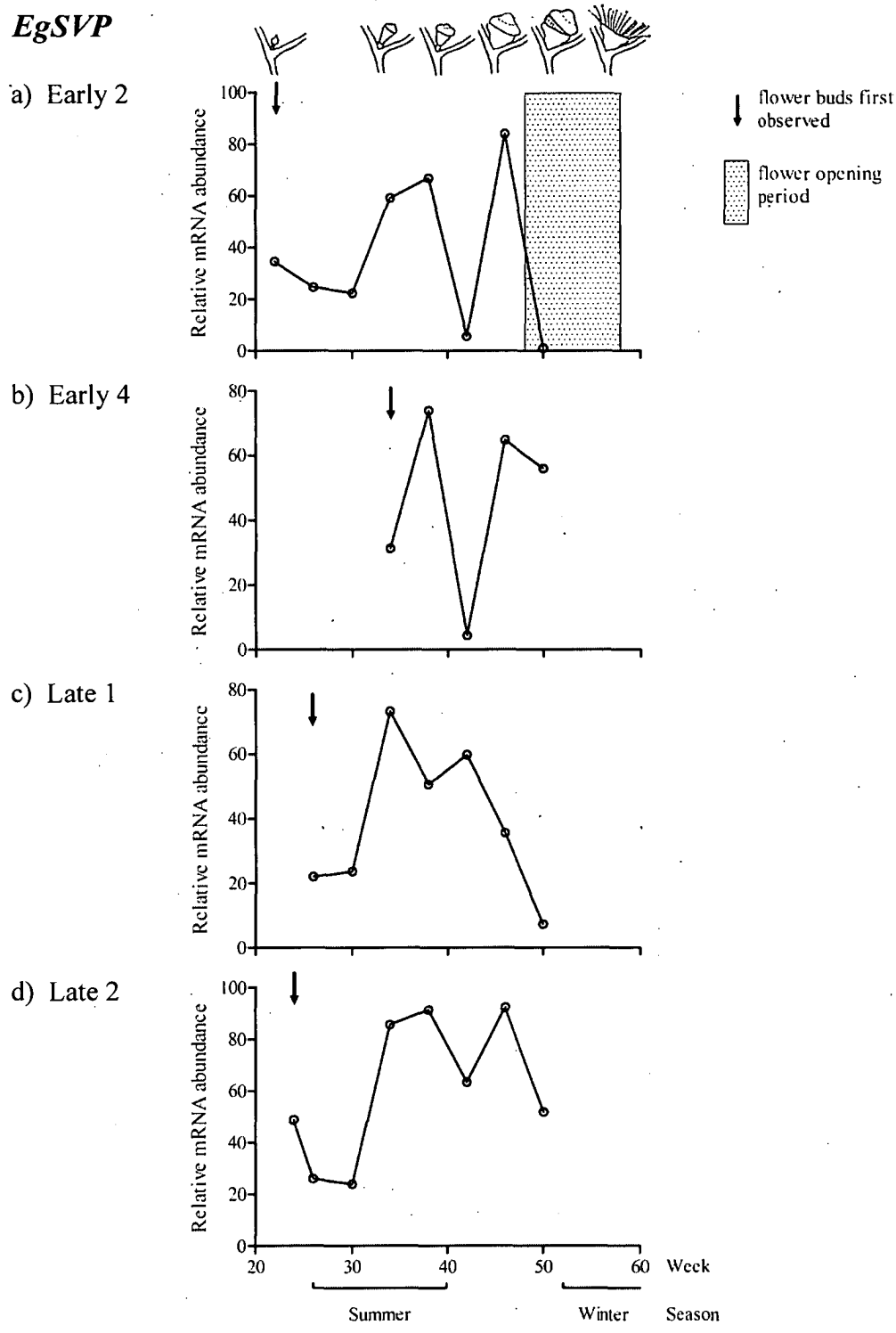


Figure 3.10. Relative *EgSVP* mRNA abundance in second year buds (H) of two ramets each of an early flower opening and late flower opening genotype of *globulus*: a) 4489-2; b) 4489-4; c) 5296-1; d) 5296-2. Timing of flower bud initiation and opening for each ramet is indicated, and representative bud morphology of 4489-2 at several sampling intervals is also shown.

3.3.4.2.2 EgTFL1

EgTFL1 transcript levels were low throughout the experiment, with an average C_T of 29.5 across all tissue types and samples (data not shown). There was no obvious seasonal pattern of *EgTFL1* expression in leaves, apices and buds of the early or late flower opening genotype (Figure 3.11, Figure 3.12). Using *EgTUBA1* instead of *Eg18S* as a reference gene also did not produce a meaningful pattern of expression of *EgTFL1* (data not shown).

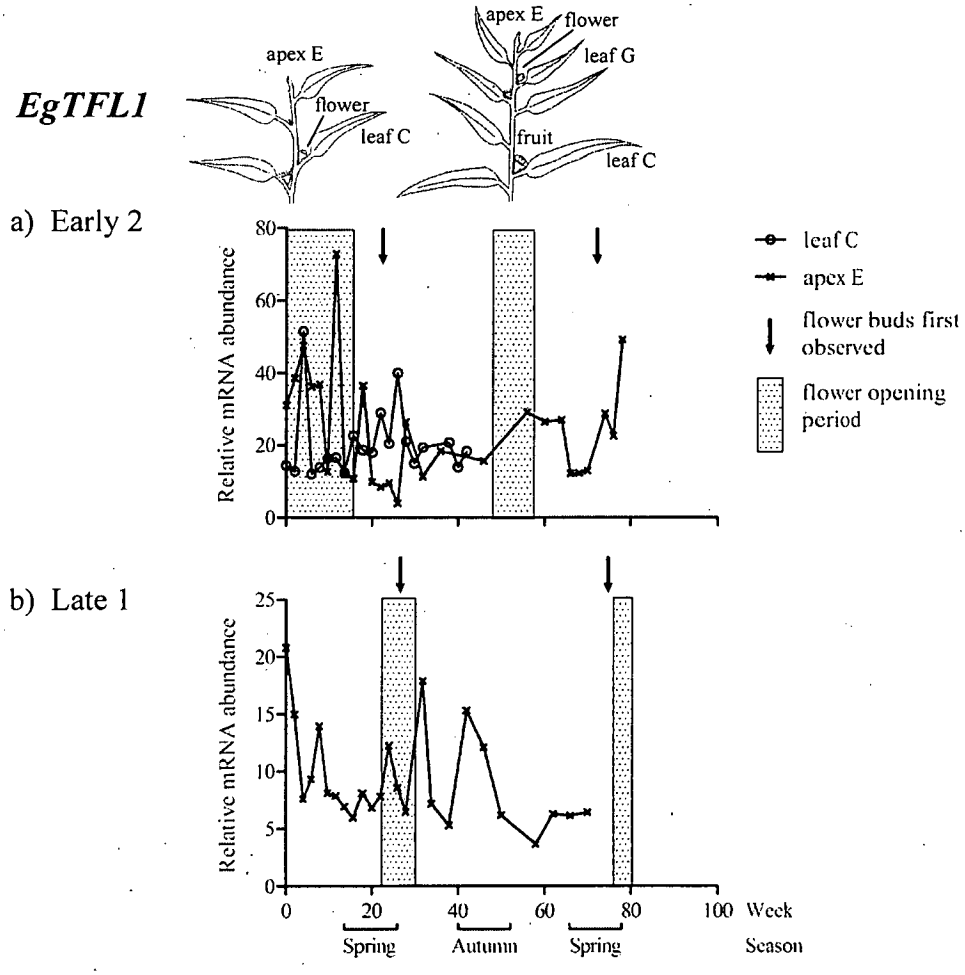


Figure 3.11. Relative *EgTFL1* mRNA abundance in leaves (C) and apices (E) of one ramet each of an early and late flower opening genotype of *globulus*: a) 4489-2; b) 5296-1. Timing of flower bud initiation and opening for each ramet is indicated, and representative shoot architecture of 4489-2 at the beginning of each year is also shown.

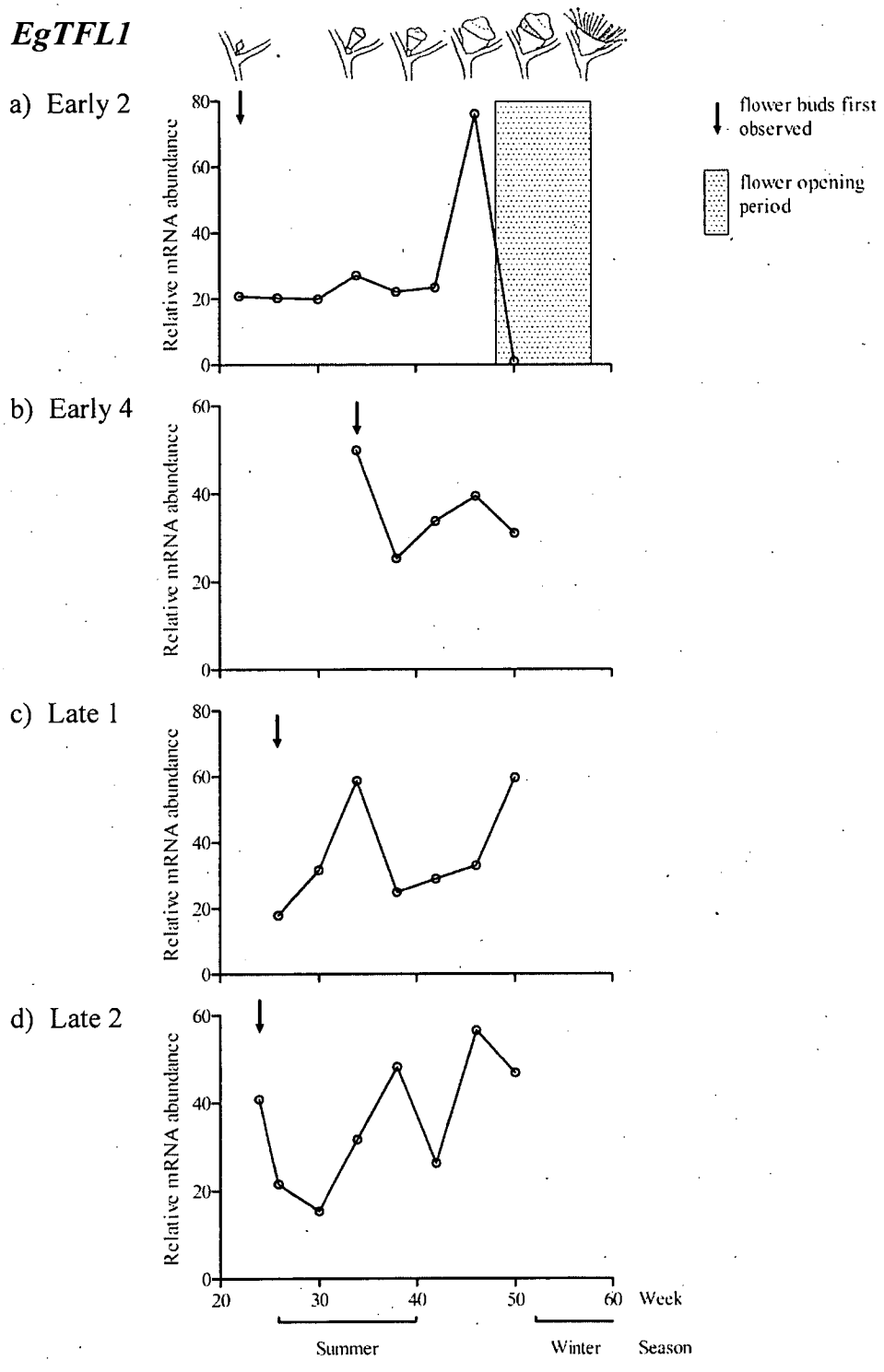


Figure 3.12. Relative *EgTFL1* mRNA abundance in second year buds (H) of two ramets each of an early flower opening and late flower opening genotype of *globulus*. a) 4489-2; b) 4489-4; c) 5296-1; d) 5296-2. Timing of flower bud initiation and opening for each ramet is indicated, and representative bud morphology of 4489-2 at several sampling intervals is also shown.

3.3.4.3 Genes conferring inflorescence identity

3.3.4.3.1 ELF1

In the Late 1 ramet, the timing of expression of *ELF1* in the apex was similar to the timing of *EgFT* expression in the leaves. It is possible that the peak in *ELF1* expression occurred after the peak in *EgFT* expression, however this could not be detected with the fortnightly (bi-weekly) harvest frequency used in this study. There was a gradual increase in the abundance of the *ELF1* transcript towards the onset of flower bud initiation in spring each year, and expression was low in late summer (Figure 3.13c). This pattern was not detected in the other ramet of this genotype (Late 2) in the first year, but in the second year there was a gradual increase in abundance of the *ELF1* transcript in the six month period before flower buds were observed (Figure 3.13d). There was no seasonal pattern of *ELF1* expression in the apex tissue of the 4489 ramets (Figure 3.13a,b).

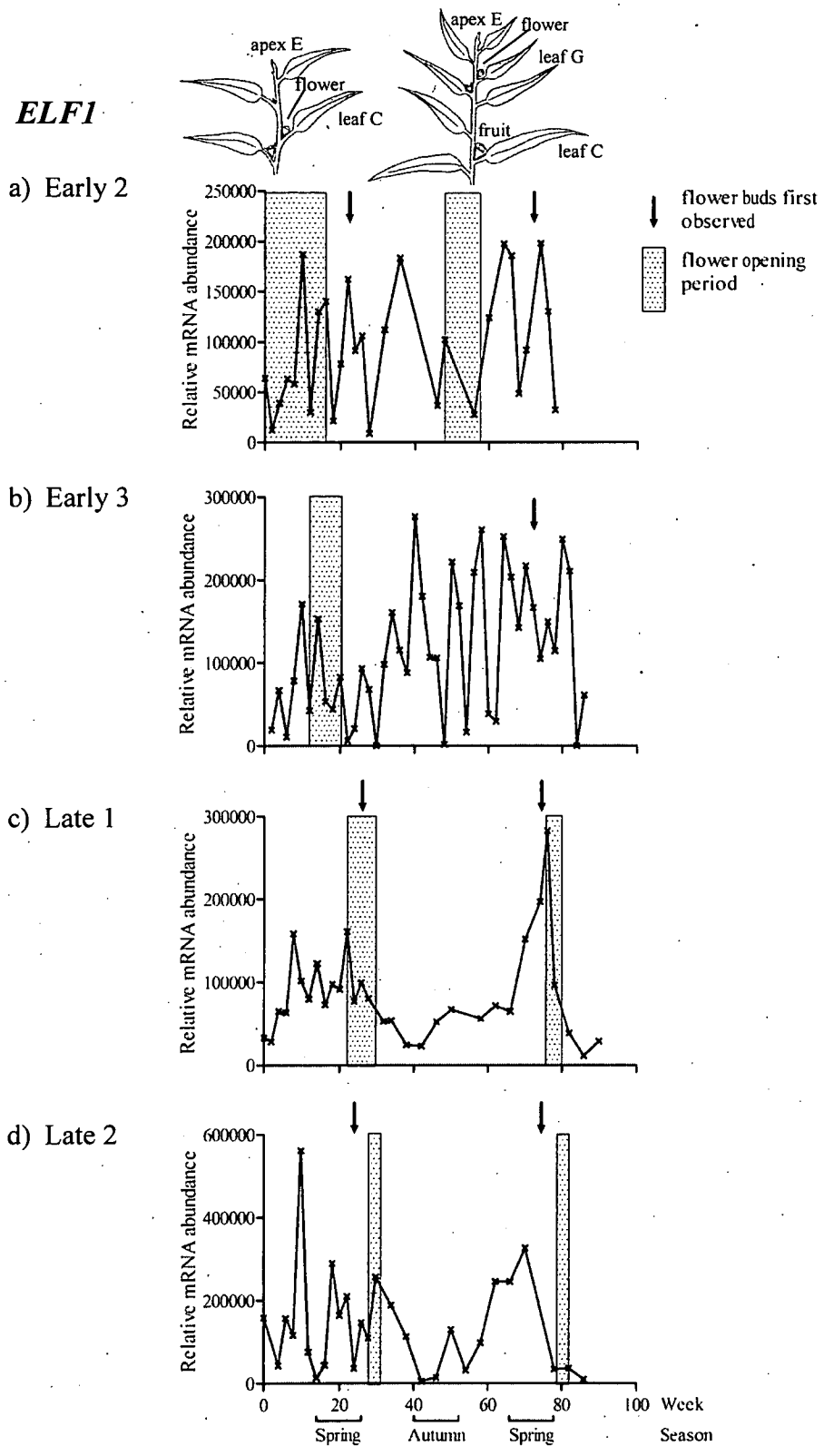


Figure 3.13. Relative *ELF1* mRNA abundance in apices (E) of two ramets each of an early and late flower opening genotype of *globulus*: a) 4489-2; b) 4489-3; c) 5296-1; d) 5296-2. Timing of flower bud initiation and opening for each ramet is indicated, and representative shoot architecture of 4489-2 at the beginning of each year is also shown.

ELF1 expression was very low in the Early 2 leaf samples across nearly a full year of harvests (data not shown), and was therefore not measured in the leaf tissue of the other three ramets.

In the second year (spring 2006 onwards), the *ELF1* transcript was very abundant in flower buds when flower buds were first visible and first able to be harvested, but *ELF1* expression in buds was low by late summer in both ramets of both genotypes (Figure 3.14), despite the difference in the rate of flower bud growth and timing of flower bud opening between the two genotypes (Table 3.2, also see Chapter 2).

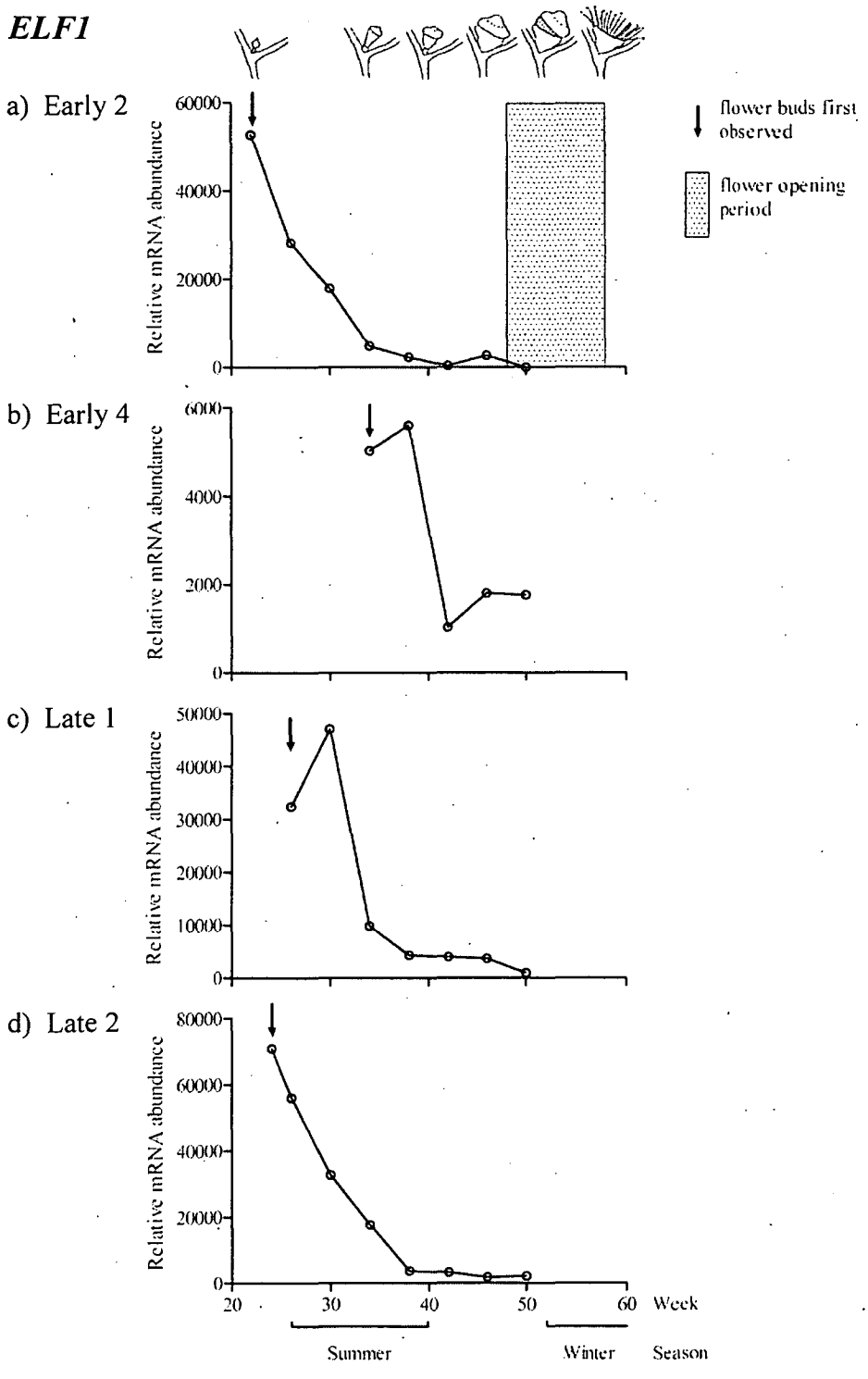


Figure 3.14. Relative *ELF1* mRNA abundance in second year buds (H) of two ramets each of an early flower opening and late flower opening genotype of *globulus*: a) 4489-2; b) 4489-4; c) 5296-1; d) 5296-2. Timing of flower bud initiation and opening for each ramet is indicated, and representative bud morphology of 4489-2 at several sampling intervals is also shown.

3.3.4.3.2 EgSOC1 and *EAP1*

There was no obvious seasonal pattern of *EgSOC1* expression in leaves and apices (Figure 3.15) nor was there a pattern of *EgSOC1* expression in bud tissue that was consistent among ramets within a single genotype (Figure 3.16). Similarly, *EAP1* expression in the leaf and apex did not follow a seasonal pattern (Figure 3.17) and the pattern of expression in flower buds was not consistent among ramets within a single genotype (Figure 3.18). In the late flower opening genotype, however, expression patterns of *EgSOC1* and *EAP1* were similar, both decreasing in abundance during flower bud development in Late 1, but both increasing in abundance during flower bud development in Late 2 (Figure 3.16, Figure 3.18).

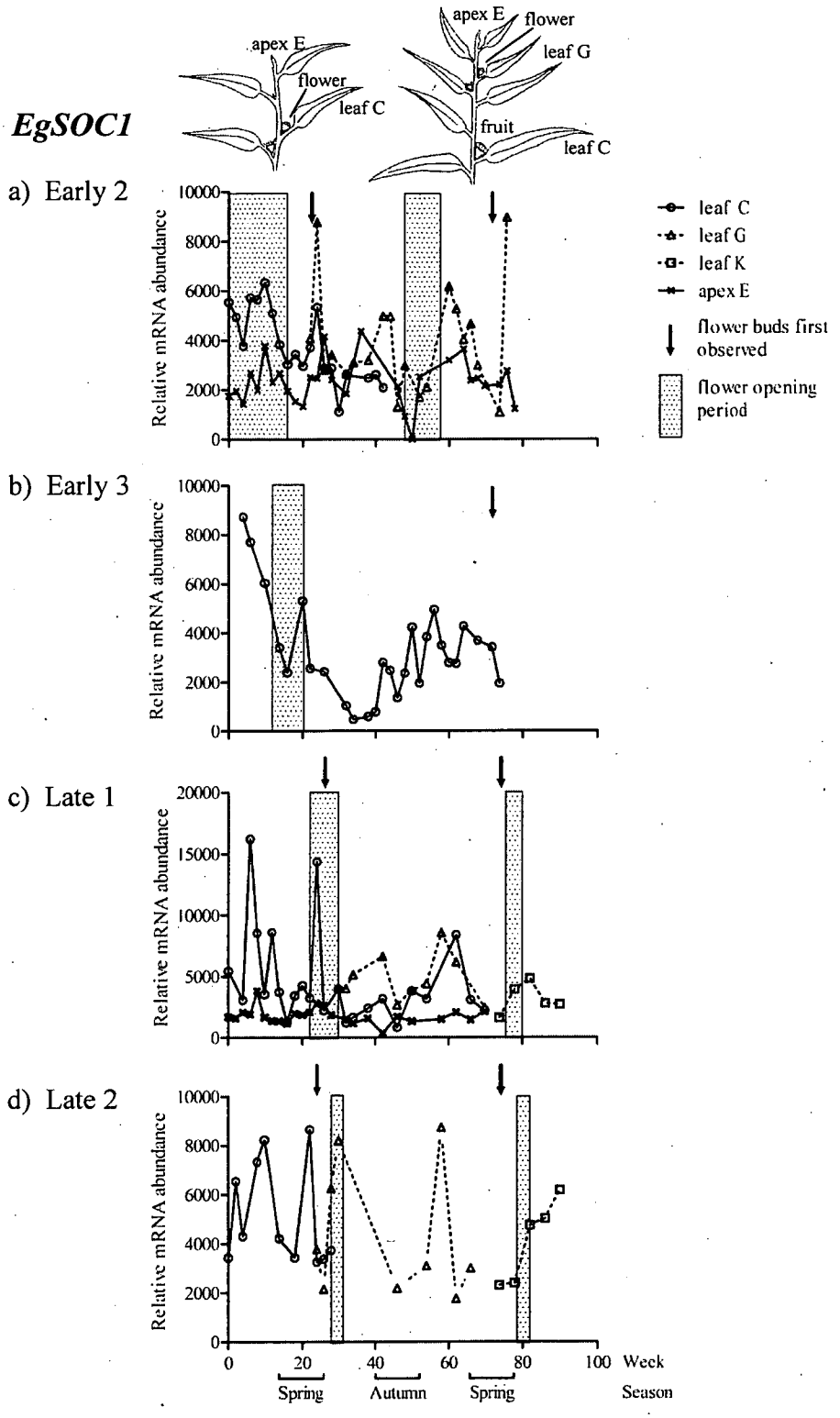


Figure 3.15. Relative *EgSOC1* mRNA abundance in leaves (C, G, K) and apices (E) of two ramets each of an early and late flower opening genotype of *globulus*: a) 4489-2; b) 4489-3; c) 5296-1; d) 5296-2. Timing of flower bud initiation and opening for each ramet is indicated, and representative shoot architecture of 4489-2 at the beginning of each year is also shown.

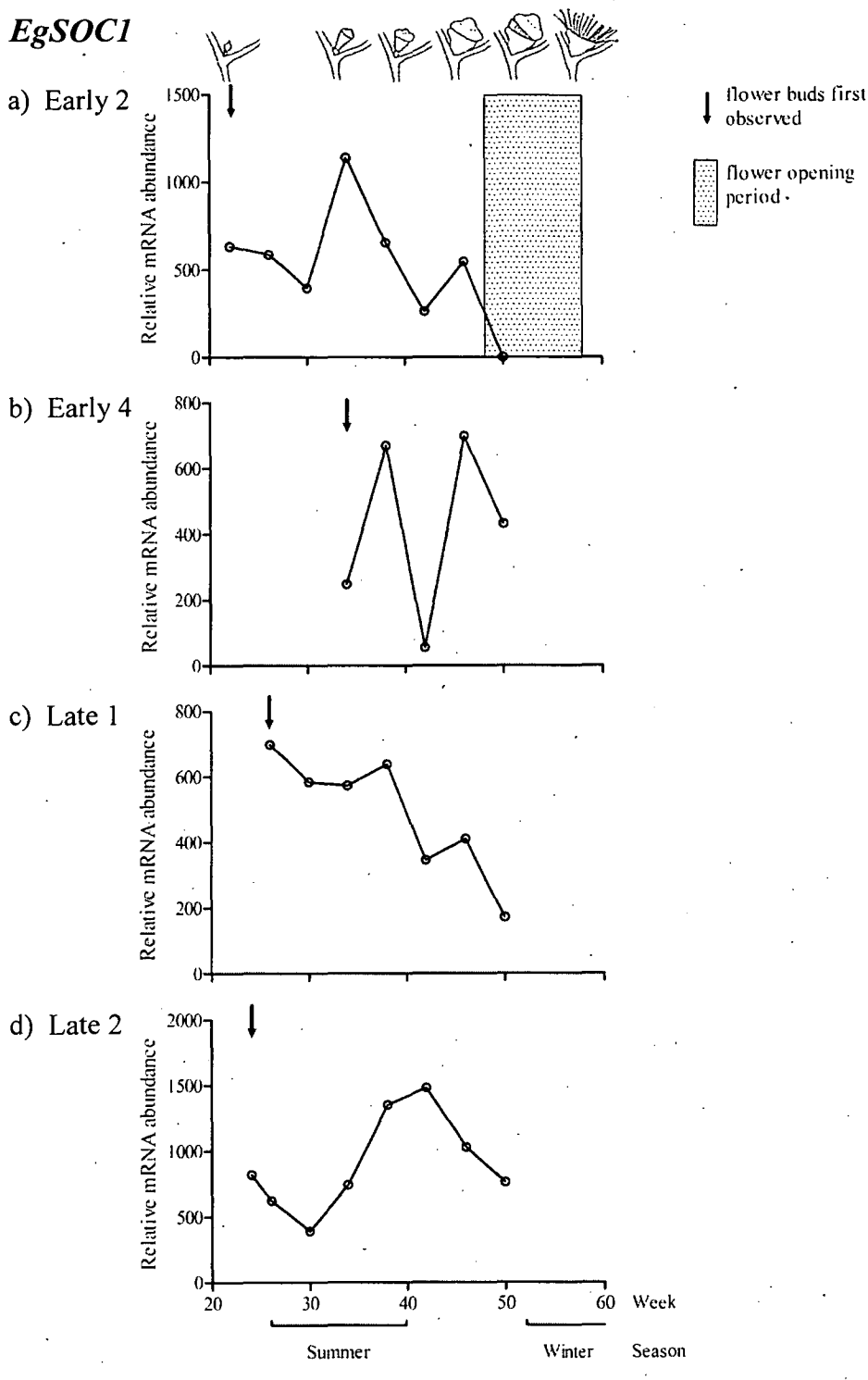


Figure 3.16. Relative *EgSOC1* mRNA abundance in second year buds (H) of two ramets each of an early flower opening and late flower opening genotype of *globulus*: a) 4489-2; b) 4489-4; c) 5296-1; d) 5296-2. Timing of flower bud initiation and opening for each ramet is indicated, and representative bud morphology of 4489-2 at several sampling intervals is also shown.

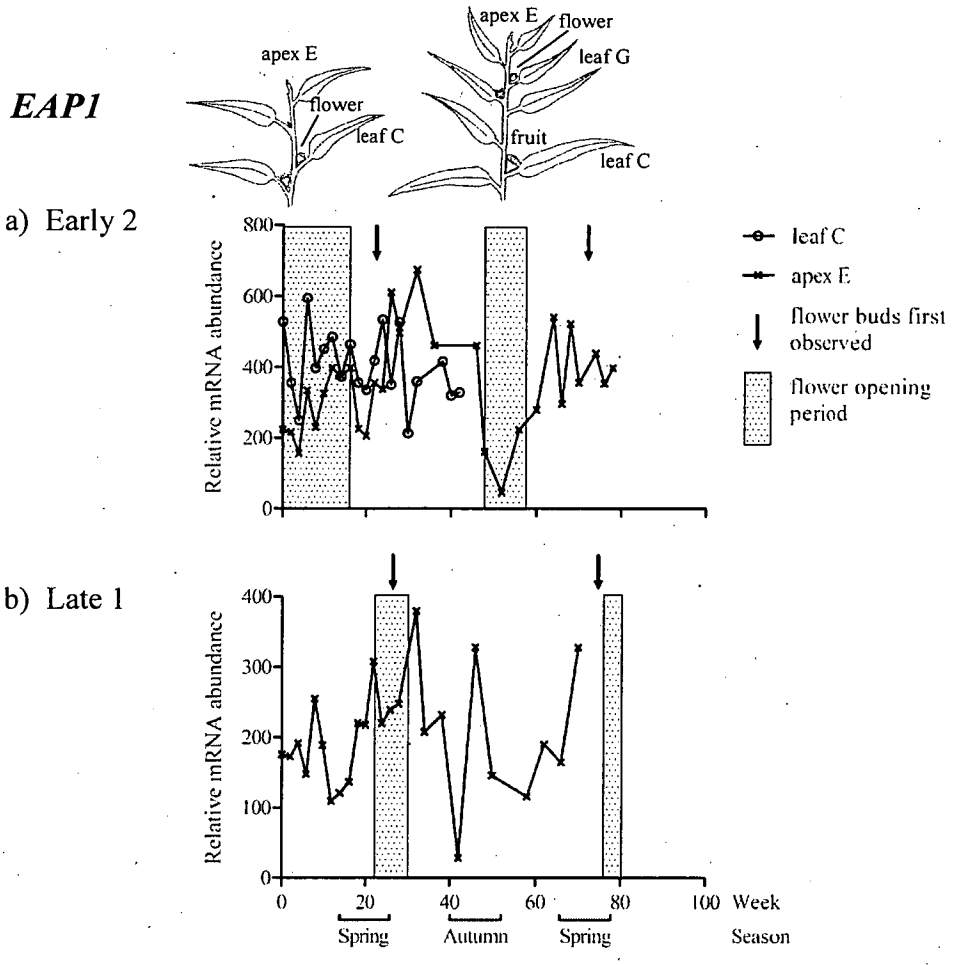


Figure 3.17. Relative *EAP1* mRNA abundance in leaves (C) and apices (E) of one ramet each of an early and late flower opening genotype of *globulus*: a) 4489-2; b) 5296-1. Timing of flower bud initiation and opening for each ramet is indicated, and representative shoot architecture of 4489-2 at the beginning of each year is also shown.

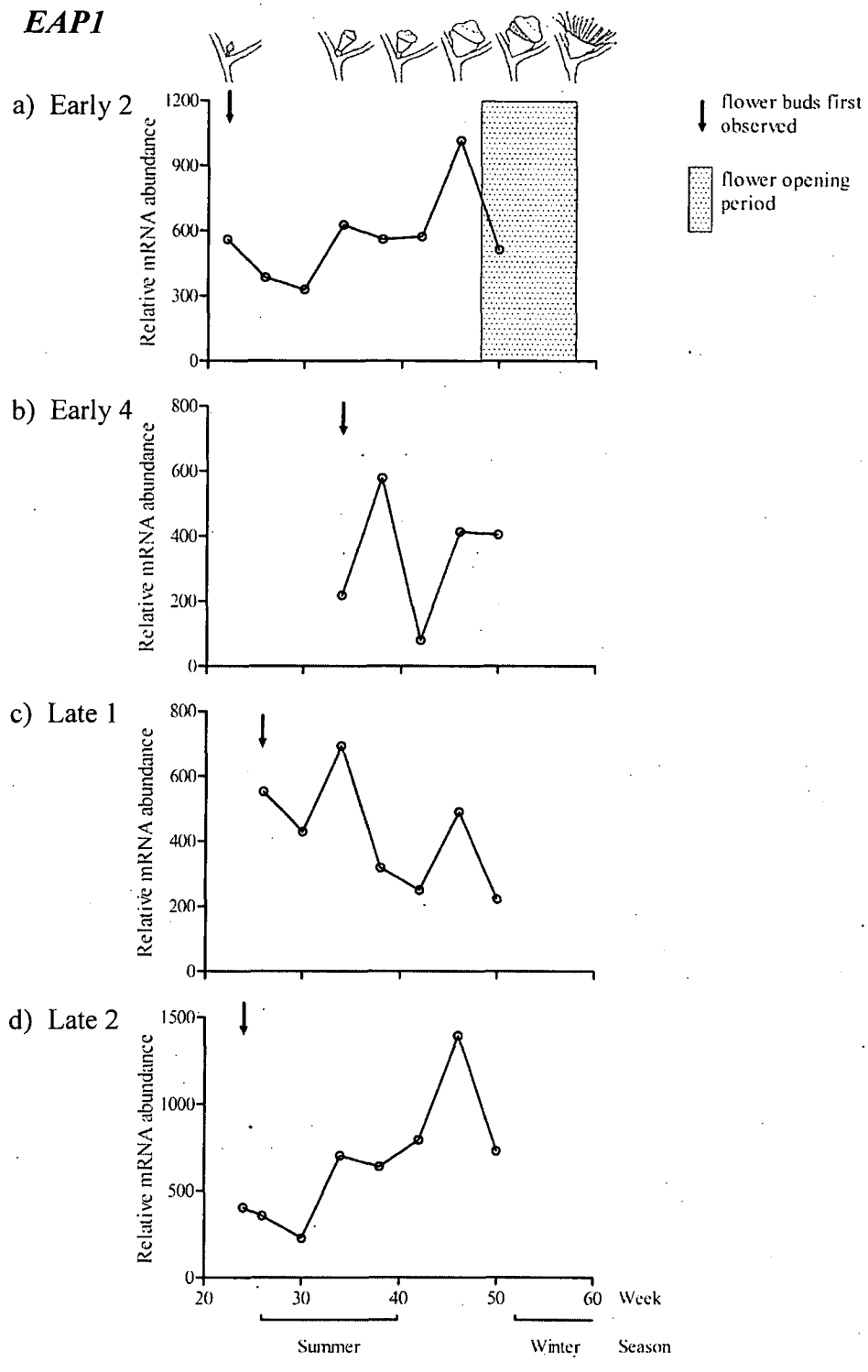


Figure 3.18. Relative *EAP1* mRNA abundance in second year buds (H) of two ramets each of an early flower opening and late flower opening genotype of *globulus*: a) 4489-2; b) 4489-4; c) 5296-1; d) 5296-2. Timing of flower bud initiation and opening for each ramet is indicated, and representative bud morphology of 4489-2 at several sampling intervals is also shown.

3.3.4.4 Genes conferring flower organ identity

As *EgFT* expression showed a temporal association with flower bud initiation but not flower opening time, the next step in the study was to examine whether genes downstream of *FT* and *LFY* in the *Arabidopsis* flowering pathway might be associated with flower bud development and opening in *globulus*.

3.3.4.4.1 EgM1, EgM2 and *EgM3*

EgM1 and *EgM3* (homologues of *SEP*) and *EgM2* (homologue of *PI*) had similar patterns of expression in Early 2 with a gradual upregulation of these three genes from mid-summer (eight weeks after flower buds were first observed) until flower bud opening in mid-autumn (Figure 3.19a; Figure 3.20a; Figure 3.21a). The troughs in this gradual upregulation coincided with the harvests at which unrepresentatively large (in terms of length and width) buds were sampled (Figure 3.22a,b compared with bud measurements for the same ramet in Figure 2.9). Despite the six month difference in flower opening time between these two genotypes (Table 3.2), the expression patterns of *EgM1*, *EgM2* and *EgM3* in Late 2 buds were similar to those of Early 2, with an initial period of eight weeks where the abundance of these transcripts was low, followed by a gradual upregulation of these genes from early summer, peaking in autumn for *EgM1* and *EgM3* (Figure 3.19d; Figure 3.21d), and in late summer for *EgM2* (Figure 3.20d). Expression profiles of *EgM1*, *EgM2* and *EgM3* for buds of Early 4 and Late 1 were noisy but also displayed a general upregulation of *EgM1*, *EgM2* and *EgM3* from when flower buds were first observed in spring until mid-autumn (Figure 3.19b,c; Figure 4.20b,c; Figure 4.21b,c). Transcripts of these three genes were less abundant in buds of Early 4 and Late 1 than in Early 2 and Late 2 (Figure 3.19, Figure 3.20, Figure 3.21).

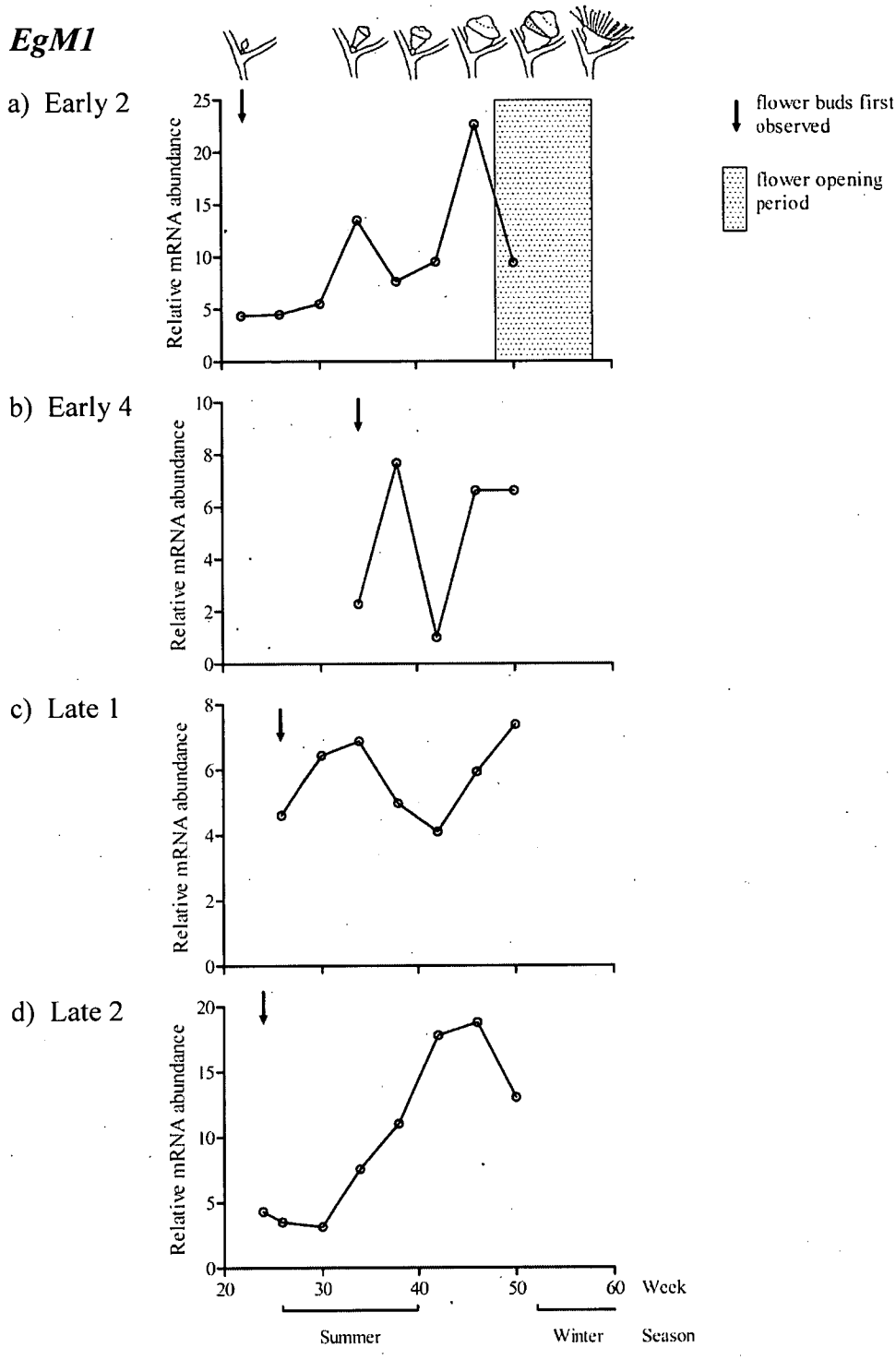


Figure 3.19. Relative *EgM1* mRNA abundance in second year buds (H) of two ramets each of an early flower opening and late flower opening genotype of *globulus*: a) 4489-2; b) 4489-4; c) 5296-1; d) 5296-2. Timing of flower bud initiation and opening for each ramet is indicated, and representative bud morphology of 4489-2 at several sampling intervals is also shown.

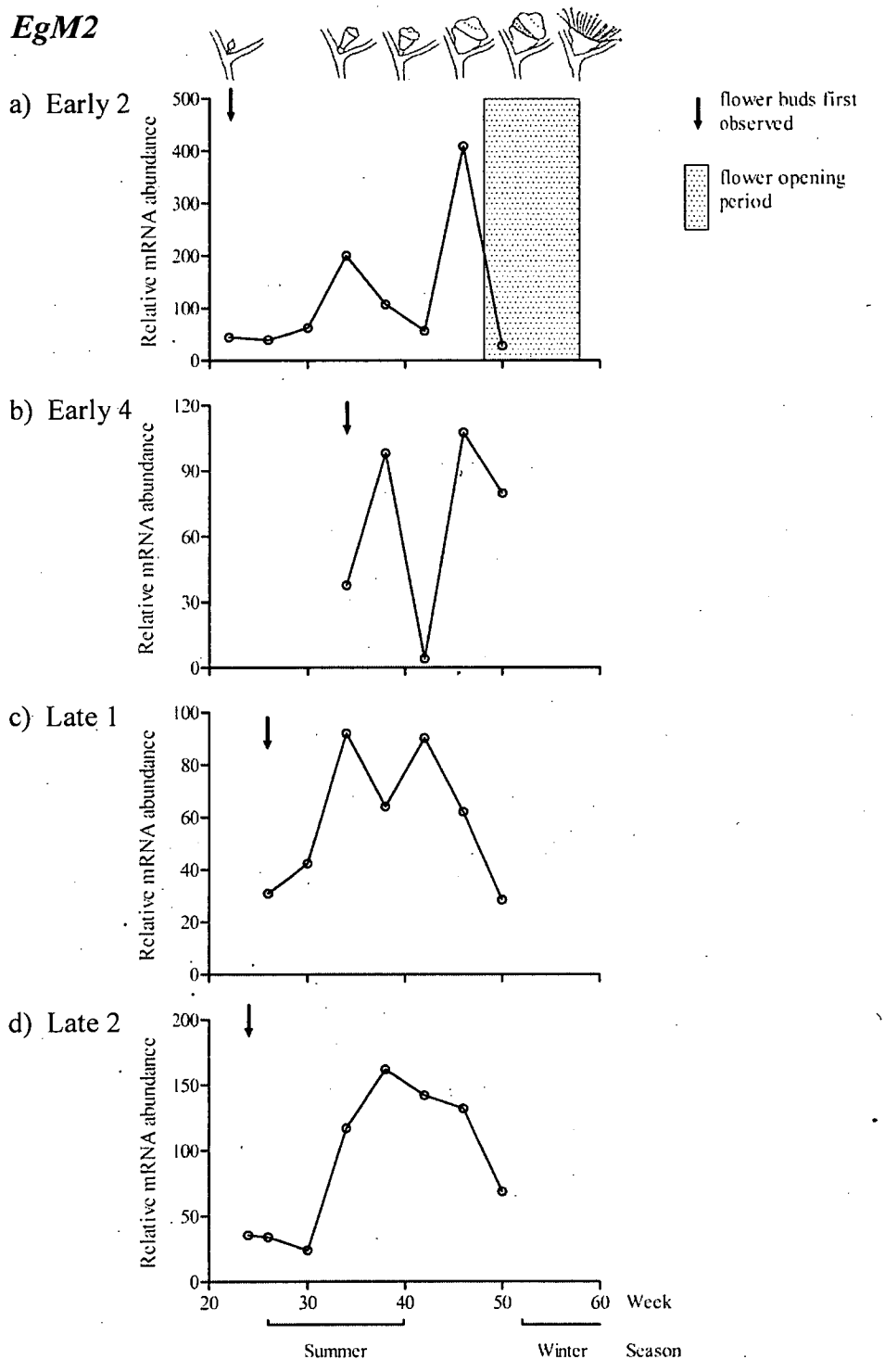


Figure 3.20. Relative *EgM2* mRNA abundance in second year buds (H) of two ramets each of an early flower opening and late flower opening genotype of *globulus*: a) 4489-2; b) 4489-4; c) 5296-1; d) 5296-2. Timing of flower bud initiation and opening for each ramet is indicated, and representative bud morphology of 4489-2 at several sampling intervals is also shown.

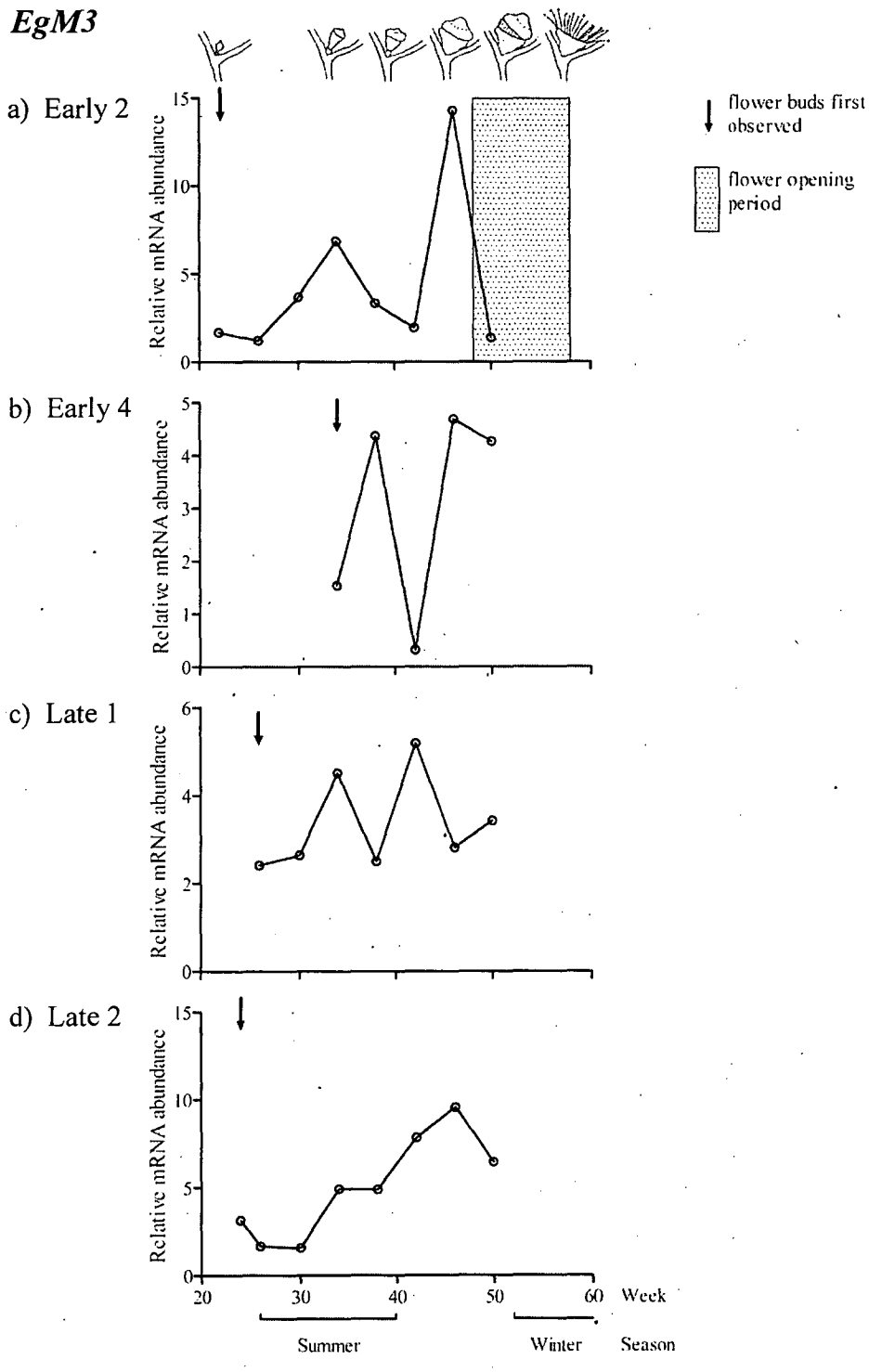


Figure 3.21. Relative *EgM3* mRNA abundance in second year buds (H) of two ramets each of an early flower opening and late flower opening genotype of *globulus*: a) 4489-2; b) 4489-4; c) 5296-1; d) 5296-2. Timing of flower bud initiation and opening for each ramet is indicated, and representative bud morphology of 4489-2 at several sampling intervals is also shown.

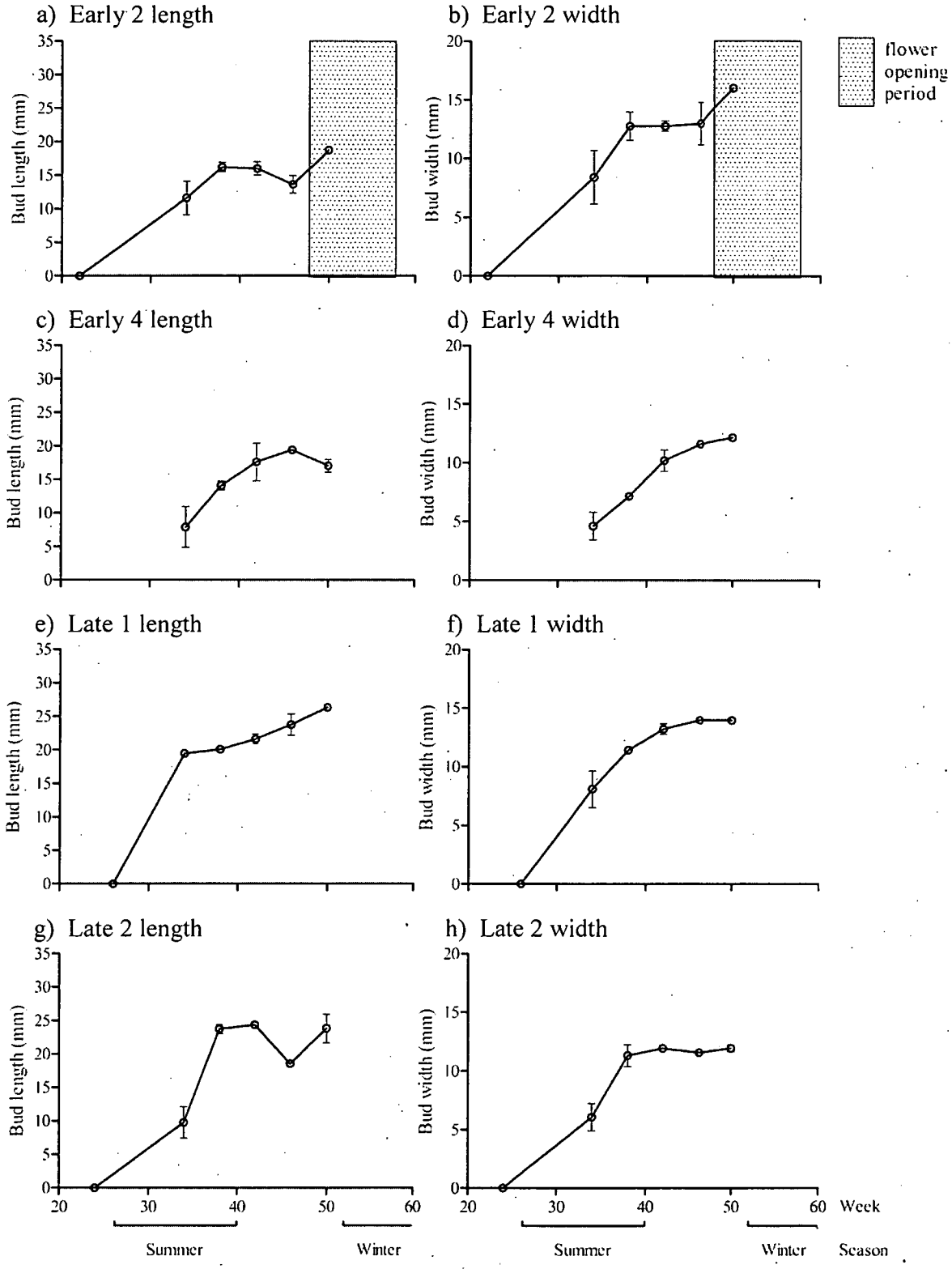


Figure 3.22. Mean length and width of buds harvested for RNA extraction. Error bars show SD. a) 4489-2 length; b) 4489-2 width; c) 4489-4 length; d) 4489-4 width; e) 5296-1 length; f) 5296-1 width; g) 5296-2 length; h) 5296-2 width.

3.4 Discussion

3.4.1 Cloning of flowering genes

The flowering genes *EgFT*, *EgTFL1*, *EgSOC1*, *ELF1*, *EAP1*, *EgSVP*, *EgM1*, *EgM2* and *EgM3* were isolated successfully from *globulus* cDNA. Apart from *LFY*, which is a single copy gene, these genes are part of gene families, but phylogenetic analysis clearly confirmed identities of each gene. *EgFT* belonged to the *FT* clade and *EgTFL1* was in the *TFL1* clade of the *FT/TFL1* gene family. The other genes were in their respective clades in the MIKC class of MADS-box genes. *EgSOC1* isolated in this study had closer homology to *AGL20* (*AtSOC1*) than the previously published *SOC1*-like sequences isolated in *E. grandis* (*EgrMADS3* and *EgrMADS4*, Watson and Brill 2004). The gDNA sequences of *EgSVP* and *EgAP1* could not be isolated, even after testing a wide range of PCR conditions and attempting to amplify each gene in two parts, suggestive of large introns in these genes.

This is the first time that a *FT* homologue has been isolated from *Eucalyptus* cDNA or gDNA and the sequence made publicly available. This may not be the only *FT*-like gene in *Eucalyptus* though, as studies of other tree species have identified more than one *FT* homologue per species. For example, at least three *FT*-like genes have been isolated in citrus (Nishikawa *et al.* 2007), four in Norway spruce (Gyllenstrand *et al.* 2007) and five in *Populus* (Bohlenius *et al.* 2006, Hsu *et al.* 2006, Igasaki *et al.* 2008). In *P. nigra var. italica*, the different members of the *FT* clade appear to have different roles; *PnFT1* and *PnFT2* appear to be associated with flower initiation while *PnFT3/4* is associated with leaf senescence (Igasaki *et al.* 2008). A blastn search of the *Eucalyptus* Genome Database (EucalyptusDB, http://eucalyptusdb.bi.up.ac.za), which includes the 4.5x checkpoint assembly of the *E grandis* genome and various EST data sets, indicated that only one *FT*-like gene was present in *Eucalyptus* so far. The true number of *FT*-like genes in *Eucalyptus* will be easier to determine once the 8x coverage assembly becomes available in 2010 (Grattapaglia and Kirst 2008) or by Southern blot.

3.4.2 Selection of reference genes

The reference genes that were tested in this study (*EgH4*, *EgTUBA1*, *Eg18S*, *EgUBI*, *EgEF1* and *EgACT11*) represented a variety of gene families and functions, and had been used as reference genes in previous expression studies in *Eucalyptus* or other forest trees. In many of these studies, however, the stability of the reference gene was assessed using RNA gel blots rather than qRT-PCR. Brunner *et al.* (2004) recommended carefully testing several housekeeping genes for studies that compare gene expression among different tissue types and developmental stages. Their linear regression analysis of housekeeping gene qRT-PCR expression levels, implemented here among samples harvested at different times of the year, provides a more sensitive method for choosing an appropriate housekeeping gene for expression studies. While the *EgTUBA1* transcript was the most stable in leaf tissue, in apices the expression of this gene was highly variable, as were *EgH4*, *EgUBI* and *EgEF1*. Use of the same reference gene to normalise target gene expression profiles for all tissue types enabled the comparison of gene expression patterns across different tissue types. The abundance of the *Eg18S* transcript was the most consistent among different tissue types and harvest dates, as evidenced by the low slope of, and close fit to, the regression line in leaf, apex and bud tissue. This validation of several reference genes provides a valuable point of reference for other studies that plan to compare gene expression among different organs and developmental stages in *Eucalyptus*.

As expected, *Eg18S* was the most abundant transcript, but this can make it an unsuitable reference gene for studying weakly expressed genes. It is more appropriate to choose a reference gene that has an expression level comparable to the genes being analysed (Bustin 2000). *TFL1* is expressed at low levels in trees such as *Metrosideros* (Myrtaceae, Sreekantan *et al.* 2004) and so a weakly expressed reference gene appeared more suitable. However, the use of either *EgTUBA1* or *Eg18S* as a reference gene still did not result in any obvious seasonal pattern of *EgTFL1* expression. Subsequently, *Eg18S* was used as a reference gene to normalise all gene expression profiles.

While the expression levels of housekeeping genes were relatively constant across different harvests of leaf samples, the same genes were less stable across different harvests of apex tissue. The pattern of expression was similar across genes for a given

harvest date, however, indicating that this was caused by variation in RNA quality or quantity, rather than instability of the internal control gene. During spring, the apex consisted of a newly expanding shoot and it is likely that harvests taken during this time included more leaf tissue, while in the winter it is possible that the sample included more stem than young leaf tissue, which could account for the variation in RNA quality or quantity. In addition, this variation could also be attributed to the variation in the developmental stage of the apex over the course of the year; the apex was dormant during autumn-winter but undergoing active growth in spring. This issue did not pose a problem in a study of the circadian clock in chestnut, where the internal control rRNA levels were similar across winter bud, leaf and stem samples (Ramos *et al.* 2005).

3.4.3 Diurnal expression patterns

The disparity in the time of day of harvesting among sampling intervals was unlikely to lead to a biased gene expression profile in this study, as there was little variation in the expression of the reference gene *Eg18S* over the 11am to 5pm period, the patterns of variation in *EgFT* and *ELF1* transcript levels over this time were inconsistent among genotypes, and the seasonal variation in *EgFT* transcripts was much greater than the fluctuation shown in the period from 11am to 5pm. The diurnal expression experiment was conducted for the period during which seasonal expression of *EgFT* was greatest, and any of the *EgFT* expression levels in the diurnal time series would have represented a peak in the seasonal time series. In *Populus deltoides*, *FT2* mRNA was most abundant in leaves at sunrise, declining to predawn levels by midday, declining slightly further by sunset (Hsu *et al.* 2006). None of the samples in this experiment were collected before midday, so, presuming that the diurnal expression pattern of *FT* in *Populus* and *Eucalyptus* is similar, this avoided the potential period of most fluctuation in the expression of *FT*.

3.4.4 Seasonal expression patterns of flowering genes

3.4.4.1 Genes associated with flower bud initiation

The expression of *EgFT* in the leaf of both early and late flower opening genotypes was associated with annual flower bud initiation, and the pattern was repeatable over a two-year period. Comparison of early and late flower opening genotypes showed no association between the expression of *EgFT* and the timing of flower bud opening though, as the expression profiles were similar despite a six month difference in the timing of flower bud opening between genotypes. Together, this indicates that the levels of *EgFT* alone are unlikely to regulate flower opening time in *globulus*, but *EgFT* could form part of the flower initiation pathway, consistent with the role proposed in other species. A similar pattern has been detected in apple, where *FT* expression peaked during the flower induction period but with only a small peak during the flower blooming (i.e. flower opening) period (Hattasch *et al.* 2008), though interestingly the tissue used in the apple study was apical meristem tissue rather than leaves. Similarly, a study of seasonal flowering gene expression patterns in Satsuma mandarin (*Citrus unshiu*) showed a seasonal increase in *CiFT* expression in the stem and leaves during floral induction (Nishikawa *et al.* 2007), however the experiment was not extended to include leaves harvested at flower opening time. The abundance of transcripts of poplar homologues of *FT* in leaves was also correlated with seasonal flower bud initiation in adult trees of *Populus deltoides* (Hsu *et al.* 2006) and *P. nigra* var. *italica* (Igasaki *et al.* 2008).

While the *EgFT* expression profile showed an overall pattern of upregulation in the lead-up to appearance of flower buds each year, the expression profile was noisy, especially in Early 2 and Late 1. In *globulus*, not all shoots initiate flowers each year, and the shoot from which some leaves were taken may have been destined to be a vegetative shoot rather than an inflorescence. These factors could explain the troughs of *EgFT* expression during the period of overall upregulation. There is also some difficulty in harvesting developmentally equivalent leaves across harvesting intervals in the case of adult forest trees, and this could also account for some of the noise in the expression profile.

The *globulus* expression data suggest that an ongoing upregulation or critical level of *EgFT* expression may be required in order to initiate flower buds, as Early 3 had a brief peak of *EgFT* expression in first year leaves but did not initiate flower buds in the first year. This is in contrast to the ongoing upregulation of *EgFT* in first year leaves of the other trees, and also in contrast to the ongoing upregulation of *EgFT* in the second year leaves of Early 3, after which flower buds were initiated. Alternatively, it may be that other factors that were absent in Early 3 in the first year but were not measured in this study play a role, such as the genes involved in the spring flush of vegetative growth.

Recent studies have shown that homologues of *FT* may have broader roles in plant development than flower initiation alone, and this could also be the case in *globulus*. In *Populus trichocarpa PtFT1* is involved in the regulation of short-day induced growth cessation (Bohlenius *et al.* 2006) and in Norway spruce (*Picea abies*), *PaFT4* expression is correlated with seasonal cycles of growth (Gyllenstrand *et al.* 2007). The possibility that the expression of *EgFT* is associated with annual growth flush rather than annual flower initiation in *globulus* cannot be discounted, as the two processes are correlated (see 2.3.1). Early 3 did not initiate flower buds in the first year of harvesting, but nor did it undergo a vegetative flush in that year. It will be necessary to analyse *EgFT* expression profiles of trees that underwent a vegetative flush but not flower bud initiation to determine with which of these two processes *EgFT* is associated.

The timing of expression of *ELF1* in the apex of the Late 1 ramet coincided with the timing of *EgFT* expression in the leaves, with a gradual upregulation of the *ELF1* transcript towards the onset of flower bud initiation each year. The pattern of expression of these two genes also coincided in apple, but there was also a peak in the apple *LFY* homologue in apices during apple flower blooming (Hattasch *et al.* 2008). However, the seasonal pattern of *ELF1* expression could not be detected in the other *globulus* ramets and requires verification.

There was no strong seasonal pattern that was repeatable among ramets of a genotype in the expression of *EgSOC1*, *EAP1* or *EgTFL1* in apex tissue, and seasonal patterns of expression of these genes in other woody perennials are contrasting. Changes in the expression of *TFL1*, *LFY* or *AP1* homologues in Satsuma mandarin (*Citrus unshiu*) stems (which included apex material) also did not correlate with floral initiation

(Nishikawa *et al.* 2007). Other studies, however, have reported activation of homologues of *SOC1* and *AP1* during flower bud initiation and/or flower development in other woody plants (Hattasch *et al.* 2008, Sreekantan *et al.* 2004, Elo *et al.* 2007). Homologues of *TFL1* have been implicated in the maintenance of the vegetative juvenile stage in citrus (Pillitteri *et al.* 2004a) and apple (Kotoda and Wada 2005, Kotoda *et al.* 2006) and the authors suggested that expression of *MdTFL1* might inhibit seasonal flower induction during apple flower opening (Hattasch *et al.* 2008). It is possible that *EgTFL1* may have a role in the transition to first flower initiation in *globulus* but not in ongoing seasonal flower initiation. The role of *ELF1*, *EgSOC1*, *EAP1* and *EgTFL1* in seasonal flower initiation in *globulus* may be elucidated by dissecting the apical tissue further, as there may be only a very small area of the apex where upregulation is important. The apices harvested in this study could have included some very young leaves with the apical meristem, and this could distort or dilute the gene expression profile.

3.4.4.2 Genes associated with flower bud development

The expression of *ELF1* was high early in bud development, declining by late summer in both genotypes. By late summer, all four flower whorls were fully developed in both genotypes, but the flower buds on the Late ramets began a period of winter dormancy whereas buds on the Early ramets continued to grow until flower opening in late April (see 2.3.3). The similar pattern of *ELF1* expression in both genotypes, despite the difference in flower bud development and opening time, implies that *ELF1* has a role in early flower bud development but not in regulating later stages of development including flower bud opening time.

The *EgM1*, *EgM2* and *EgM3* genes, which are homologous to *SEP3*, *PI* and *SEP1*, respectively, act downstream of *LFY* to control the identity of flower organs. These genes were gradually upregulated in Early 2, starting when flower buds were first observed and continuing until flower bud opening time. The noise in the profile can be accounted for by the unrepresentative buds that were harvested at some harvest intervals. In other ramets though, the pattern was less clear. Southerton *et al.* (1998a) showed that expression of *EgM2* was strongest in eucalypt petals (operculum) and stamens, consistent with its homology to class B genes of the ABC model of flower development,

and *EgM3* was strongly expressed in the receptacle, which surrounds the carpels and holds the sepals, petals and stamens. It is likely that the expression patterns of these genes were distorted in this study, firstly, by harvesting entire buds rather than dissecting individual flower parts; and secondly, because of the difficulty in destructive harvesting of equivalent buds over time on a tree bearing multiple flowering branches that did not necessarily initiate flowers simultaneously.

SVP expression has been associated with winter dormancy of lateral vegetative buds in apricot (Yamane *et al.* 2008). The peak of *EgSVP* expression in the flower buds of *globulus* Late ramets also coincided with the onset of their dormancy in late February. There was also an association between *EgSVP* expression and vegetative dormancy in the apex, but this association was not observed in all trees and must be confirmed with further expression profiling in extra ramets. The potential role of *EgSVP* in promoting winter dormancy in flower buds and shoot apices is worth exploring further, as the timing of release from dormancy may be a direct driver of flower opening time (Chapter 3).

3.4.4.3 Potential genes associated with flower bud opening

While the seasonal peaks in *EgFT* expression are associated with *globulus* flower bud initiation, the genes associated with seasonal flower bud opening are still unknown. As flower bud opening time is a key trait of economic importance (see 3.1.3) the genes that are potentially associated with this process are worth pursuing. Seasonal flower bud opening occurs once all parts of the flower bud are fully developed, and the inner operculum (fused petals) abscission zone could be developmentally equivalent to the petal abscission zone in other species. The signals involved in flower organ abscission have been well-documented in *Arabidopsis* and, while the process in trees may not be directly comparable, *Arabidopsis* provides a useful starting point for investigating the possible genes responsible for the observed variation in flower opening time in *globulus*.

Timing of abscission of flowers and flower parts is controlled by the hormone ethylene in many plant species (Dugardeyn and Van der Straeten 2008). Variation in the expression of 1-aminocyclopropane-1-carboxylic-acid (ACC) synthase and ACC oxidase, key rate-limiting genes of the ethylene synthesis pathway (Johnson and Ecker

1998), could contribute to the variation in flower opening time among genotypes of *globulus*. The *Eucalyptus* sequences of ACC synthase and ACC oxidase are not available, but these genes have been isolated in poplar (Vahala *et al.* 2003, Ingvarsson unpubl.) and degenerate primers could be used to isolate the *Eucalyptus* homologues and analyse the expression patterns of these genes during flower bud development. Ethylene inhibitors have been used to extend vase life of cut flowers because it delays petal abscission (e.g. in roses, Sane *et al.* 2007). There have been no studies using ethylene inhibitors to delay petal abscission (i.e. flower opening) in *Eucalyptus*, but ethylene inhibitors have been used to prevent leaf senescence and extend the vase life of cut foliage of eucalypts (Ferrante *et al.* 2002).

In *Arabidopsis*, the signalling pathway that controls abscission also includes genes such as *INFLORESCENCE DEFICIENT IN ABSCISSION* (*IDA*), *HAESA* (*HAE*) and *HAESA-LIKE 2* (*HSL2*), *MITOGEN-ACTIVATED PROTEIN KINASE KINASE 4* (*MKK4*) and *MKK5* (reviewed in Cho *et al.* 2008). Homologues of these genes could be further candidates to account for the variation in *globulus* flower opening time.

The variation in flower opening time among genotypes of *globulus* might also be attributable to variation in the development of the abscission zone of the operculum, due to mutations in genes directly involved in the development of the abscission zone. A number of genes have been implicated in abscission zone development in plants (reviewed in Taylor and Whitelaw 2001). For example, the overexpression of *AGL15* causes abscission zone cells to be maintained in the juvenile state in *Arabidopsis*, delaying abscission of petals (Fernandez *et al.* 2000). Genes like this that are involved in abscission zone development could, therefore, also be worth investigating in *globulus*.

3.4.5 Future directions in eucalypt flowering studies

The *Eucalyptus grandis* genome sequence, which will be released into the public domain in 2010 (Grattapaglia and Kirst 2008), will be a valuable resource for genomic and biotechnology research in *Eucalyptus*. It will be relatively easy to identify *Eucalyptus* homologues of *Arabidopsis* genes for expression analyses, and also identify cases where there is no *Eucalyptus* counterpart. Although *FLC* has not been identified outside the

Brassicaceae so far, a *Eucalyptus* homologue of *FLC* has been identified in the FORESTS database (Dornelas and Rodriguez 2005a), but the sequence is not publicly available. With the release of the *E. grandis* genome, the role of this gene in regulating floral initiation in *Eucalyptus* could be elucidated.

Integrating the *Eucalyptus* genetic map and the *E. grandis* genome sequence will also assist in QTL studies, to link natural variation in flower initiation and flower opening with functional genomics. QTL analysis of a large F_2 family with grandparents that differed in time to first flowering and vegetative phase change are currently being undertaken at UTAS, and may identify QTLs associated with these traits. If the confidence intervals around these QTLs are not too large and the molecular markers can be found in the *E. grandis* genome sequence, positional candidate genes may be identified. A precocious flowering mutant of *E. grandis* has been identified (Missiaggia *et al.* 2005) and early flowering can been induced in *E. occidentalis* (Southerton 2007), and these may also be useful tools for advancing molecular research into flowering of eucalypts.

CONCLUSIONS

This study evaluated the morphological and molecular diversity and evolutionary relationships among populations of the *E. globulus* species complex, the variation in flowering time among the main provenances of *globulus* used in breeding programs, isolated the *globulus* homologues of genes that are known to be associated with flowering in the model annual plant *Arabidopsis* and investigated their patterns of expression.

Morphological and molecular analyses of populations of the *E. globulus* complex showed that the evolutionary processes of drift, selection and hybridisation have all had a role in shaping the patterns of variation in *E. globulus*. While the evidence suggests that much of the intermediate zone is differentiated by selection and isolation by distance rather than recent hybridisation, there is some evidence of secondary contact (hybridisation) in parts of the *E. globulus* range. In some geographically or ecologically isolated populations, genetic drift and inbreeding have resulted in reduced genetic diversity and marked molecular divergence from the rest of the gene pool. This has implications for conservation planning, and information from this study has already been used in an environmental impact assessment of a proposed windfarm site. Overall, geography, rather than morphology, proved to be the best predictor of genetic affinities between populations and this information can be used to guide seed collections for reforestation and breeding programs.

The morphological and molecular data indicate that *bicostata*, *maidenii* and *globulus* are clear taxonomic entities as they are clearly differentiated and their cores are fairly uniform. However the subspecies status of *pseudoglobulus* is questionable as it is the least differentiated taxon, it has a limited distribution and it has high morphological variability in its core distribution. It appears to be part of a widespread lineage of variable morphology that extends across Victoria on the southern side of the Great Dividing Range. As there is continuous morphological and molecular variation between the taxa, and the intergrade populations occupy a large proportion of the gene pool, *bicostata*, *maidenii* and *globulus* are best treated as subspecies of *E. globulus* rather than separate species of a species complex.

The patterns of genetic diversity in the *E. globulus* species complex observed in this study suggest that the origin of the complex is in the SE Forests region, near the border of NSW and Victoria. It is proposed that *E. globulus* dispersed westward on the coastal side of the Great Dividing Range, before migrating from western Victoria to the northern side of the Great Dividing Range and into Tasmania via a west coast route through King Island. The Great Dividing Range appears to have provided a strong barrier to dispersal, while the Bass Strait and the Murray Darling Depression have provided incomplete barriers, the strength of which varied with changes in climate, glacial activity and sea levels associated with glacial cycles.

While the differentiation in microsatellite profiles across the range of *E. globulus* could be due to reduced gene flow resulting from geographical isolation, the genetically based differences in annual flower opening time identified in this study could also play a role. There were significant differences among races of *globulus* in the timing of flower bud opening, with Furneaux Group and East Coast South races flowering earliest each year, and Strzelecki Ranges and Otway Ranges races flowering latest, and this trait was highly heritable at least in the broad sense. There was, however, also variation from year to year in the onset and duration of flowering, so the barrier to gene flow between the early and late flowering races is unstable and there was evidence that, overall, heat sum is a major driver of this variation. Understanding the variation in flowering time has practical implications not only for future management of seed orchards, but also in assessing the risk of gene flow from planted *globulus* into native stands, and also in managing native stands for the benefit of the swift parrot that feeds on the nectar of *globulus* flowers. Information on the genotype or provenance deployed and the climate of the local area could eventually be used in a predictive model of flower opening time in a seed orchard, plantation or native stand.

While there was a large amount of genetically-based variation in flower opening time, flower bud initiation occurred nearly simultaneously in all clones each year. It appears that flower bud initiation could be a photoperiod-mediated process. Measurements taken on developing flower buds showed that the differences in flower opening time were due mainly to slower bud development and a period of bud dormancy in late flowering genotypes, while the buds of early flowering genotypes did not undergo a

dormant period. This implies that a model combining heat sum and chilling units would be worth testing. It may be that there is no chilling requirement in early genotypes, while in later genotypes a chilling period is required, and the heat sum after this chilling period determines flower opening time.

The next part of this study was to identify the genes that may be controlling the variation in flowering traits, from a list of candidate genes that are known to be associated with flowering processes in the model annual plant *Arabidopsis*, or other forest trees. *EgFT*, *EgTFL1*, *EgSOC1*, *ELF1*, *EAP1*, *EgSVP*, *EgM1*, *EgM2*, *EgM3* were isolated successfully from *globulus* cDNA. *EgSOC1* isolated in this study had closer homology to *AGL20* (*AtSOC1*) than the previously published *SOC1*-like sequences isolated in *E. grandis* (*EgrMADS3* and *EgrMADS4*, Watson and Brill 2004). This is the first time that a *FT* homologue has been isolated from *Eucalyptus* cDNA or gDNA and the sequence made publicly available.

Before gene expression studies can be undertaken, it is necessary to determine which reference gene is the most appropriate for the tissue types and developmental stages to be sampled. The validation of several reference genes in this thesis will provide a valuable point of reference for other projects that plan to compare gene expression among different organs and developmental stages in *Eucalyptus*. In this study, the abundance of the *Eg18S* transcript was the most consistent among different tissue types and harvest dates.

Analysis of expression levels of the *globulus* homologues of the flowering genes over a two-year period showed that *EgFT* expression in the leaf of both early and late flower opening genotypes was associated with annual flower bud initiation. Comparison of early and late flower opening genotypes, however, showed no association between the expression of *EgFT* and the timing of flower bud opening as the expression profiles were similar despite a six month difference in the timing of flower bud opening between genotypes. Together, these data indicate that the levels of *EgFT* alone are unlikely to regulate flower opening time in *globulus*, but *EgFT* could form part of the flower initiation pathway, consistent with the role proposed in other species. There was also some evidence that *ELF1* expression in the apex was associated with flower bud

initiation, and this gene also had a role in early flower bud development but not in regulating later stages of development including flower bud opening time.

While the seasonal peaks in *EgFT* expression are associated with *globulus* flower bud initiation, the genes associated with seasonal flower bud opening are still unknown. As flower bud opening time is a key trait of economic importance, the genes associated with this process are worth pursuing; genes involved in flower organ abscission and ethylene biosynthesis in *Arabidopsis* are good candidates. The *Eucalyptus grandis* genome sequence that will be released into the public domain in 2010 (Grattapaglia and Kirst 2008) will be a valuable resource for the identification of these and other genes that may be responsible for the observed variation in flower opening time in *globulus*.

With clear evidence of global climate change, there is considerable concern over whether trees have the ability to adapt quickly to rapidly changing environments. The limited seed dispersal mechanisms in eucalypts (Potts and Wiltshire 1997) will impede their ability to migrate to more suitable climates, and adaptation will be imperative. Variation in quantitative traits will influence the adaptive potential of populations probably more than variation in neutral markers; however, the variation across the *E. globulus* species complex in genes affecting quantitative traits, such as flowering time, remains to be established. The availability of the *E. grandis* genome sequence will lead to the development of tools to enable population genomics studies to proceed. Such studies will also help to resolve the role of selection relative to other evolutionary processes in shaping the observed complex patterns of variation in *E. globulus*.

APPENDICES

Appendix 1. Subset of core samples used for the AMOVA of the *E. globulus* species complex.

Taxon (Code)	Region Code	Region Name	Locality Code	Locality Name	*n*
bicostata(B)	3	Hume	3	Tallarook	9
		Hume	4	Euroa	12
		Hume	5	Eildon	9
	4	NE Victoria	6	Mt Granya	8
		NE Victoria	7	Shelley	8
		NE Victoria	8	Nariel	8
		NE Victoria	9	Mt Unicorn	8
	5	Canberra	10	Bungongo	13
		Canberra	11	Burrinjuck	7
bicostata total					82
maidenii (M)	8	Araluen	14	Araluen	16
		Araluen	15	Currowan-Monga	23
	9	Wadbilliga	16	Wadbilliga North	14
		Wadbilliga	17	Belowra Rd	10
		Wadbilliga	18	Wadbilliga South	20
		Wadbilliga	19	Murrabrine	20
		Wadbilliga	20	Kooraban NP	6
		Wadbilliga	21	Mt Dromedary (NSW)	13
		Wadbilliga	22	Mumbulla SF	20
		Wadbilliga	23	Brown Mountain	12
	10	SE Forests	24	Tantawangalow	8
		SE Forests	25	Big Jack - Rocky Hall	15
		SE Forests	26	Yurammie	16
		SE Forests	27	Mt Imlay	8
maidenii total					201
pseudoglobulus (P)	13	Lakes Entrance	38	Ostlers Rd	9
		Lakes Entrance	39	Lake Tyers	14
		Lakes Entrance	40	Lakes Entrance	11
		Lakes Entrance	41	Metung-Lakes Entrance	6
		Lakes Entrance	42	Metung	8
pseudoglobulus total					48
globulus	28	Recherche Bay	77	Recherche Bay	27
	29	Southern Tasmania	78	South Bruny Island	8
		Southern Tasmania	79	Dover	7
		Southern Tasmania	80	Geeveston	10
	30	SE Tasmania	81	Tinderbox	9
		SE Tasmania	83	Maria Island South	30
		SE Tasmania	84	Maria Island North	6
		SE Tasmania	85	Triabunna	8
	31	Dromedary Tas	86	Mt Dromedary (Tas)	19
		Dromedary Tas	87	Platform Peak	17
	32	NE Tasmania	88	Jericho	6
		NE Tasmania	89	Mayfield South	6
		NE Tasmania	90	Mayfield North	8
		NE Tasmania	91	Cape Tourville Dwarf	20
		NE Tasmania	92	Cape Tourville Tall	38
		NE Tasmania	93	Pepper Hill	8
		NE Tasmania	94	Humbug Hill	9
globulus total					236
E. globulus Grand Total					567

Appendix 2. Subset of samples used for analysis of genetic affinities of intergrade populations of the *E. globulus* species complex.

Taxon (Code)	Region Code	Region Name	Locality Code	Locality Name	*n*
bicostata (B)					
	2	Mt Cole	2	Mt Cole	19
	3	Hume	3	Tallarook	9
		Hume	4	Euroa	12
		Hume	5	Eildon	9
	4	NE Victoria	6	Mt Granya	8
		NE Victoria	7	Shelley	8
		NE Victoria	8	Nariel	8
		NE Victoria	9	Mt Unicorn	8
	5	Canberra	10	Bungongo	13
		Canberra	11	Burrinjuck	7
bicostata total					101
maidenii (M)					
	8	Araluen	14	Araluen	16
		Araluen	15	Currowan-Monga	23
	9	Wadbilliga	16	Wadbilliga North	14
		Wadbilliga	17	Belowra Rd	10
		Wadbilliga	18	Wadbilliga South	20
		Wadbilliga	19	Murrabrine	20
		Wadbilliga	20	Kooraban NP	6
		Wadbilliga	21	Mt Dromedary (NSW)	13
		Wadbilliga	22	Mumbulla SF	20
		Wadbilliga	23	Brown Mountain	12
	10	SE Forests	24	Tantawangalow	8
		SE Forests	25	Big Jack - Rocky Hall	15
		SE Forests	26	Yurammie	16
		SE Forests	27	Mt Imlay	8
		SE Forests	28	north Cann Valley Highway	7
		SE Forests	29	Wroxham	8
maidenii total					216
maidenii - *pseudoglobulus* intergrade (m-p)					
	11	Alfred-Nadgee	30	Nadgee	18
		Alfred-Nadgee	31	Maramingo Creek	10
		Alfred-Nadgee	32	Alfred NP	9
		Alfred-Nadgee	33	mid Cann Valley Highway	6
		Alfred-Nadgee	34	south Cann Valley Highway	12
maidenii - *pseudoglobulus* intergrade total					55
pseudoglobulus (P)					
	12	Mt Cann	35	Mt Cann	10
		Mt Cann	36	Wiebens Hill	8
		Mt Cann	37	Murrungowar	8
	13	Lakes Entrance	38	Ostlers Rd	9
		Lakes Entrance	39	Lake Tyers	14
		Lakes Entrance	40	Lakes Entrance	11
		Lakes Entrance	41	Metung-Lakes Entrance	6
		Lakes Entrance	42	Metung	8
	14	Lerderderg Gorge	43	Lerderderg Gorge	20
pseudoglobulus total					94
pseudoglobulus-bicostata-maidenii intergrade (b-m-p)					
	15	Buchan	44	Stoney Creek	10
		Buchan	45	Cutts Creek Rd	8
		Buchan	46	Buchan	11
		Buchan	47	Gelantipy	8
pseudoglobulus-bicostata-maidenii intergrade total					37

Taxon (Code)	Region Code	Region Name	Locality Code	Locality Name	*n*
bicostata-pseudoglobulus intergrade (b-p)					
	16	Omeo	48	Beloka Rd	7
		Omeo	49	Omeo	8
		Omeo	50	Trapyard Gap	7
	17	Mitchell River	51	Cobbannah	10
		Mitchell River	52	Peel Gap	10
		Mitchell River	53	Heyfield	4
bicostata-pseudoglobulus intergrade total					46
globulus intergrade Victoria (g)					
	18	Strzelecki Ranges	54	Bowden-Carrajung	2
		Strzelecki Ranges	55	Jeeralang	23
	19	Eastern Otways	56	Lorne	14
		Eastern Otways	57	Jamieson Creek	8
		Eastern Otways	58	Cape Patton	7
	20	Western Otways	59	Parker Spur	8
		Western Otways	60	Cannan Spur	8
		Western Otways	61	Otway State Forest	8
		Western Otways	62	Lavers Hill	8
globulus intergrade total					86
globulus Victoria (G)					
	24	South Gippsland	71	Alberton West-Welshpool-Hedley	14
		South Gippsland	72	Toora	8
		South Gippsland	73	Fish Creek	6
	25	Tidal River	74	Tidal River	31
	26	Wilson's Promontory	75	Wilson's Promontory Lighthouse	31
	27	Phillip Island	76	Phillip Island	24
globulus total					114
E. globulus Grand Total					749

Appendix 3. *A priori* groupings for comparison of morphological and molecular affinities of mainland Australia intergrade *E. globulus* samples.

	Region		Locality			
Taxon (Code)	Code	Name	Code	Name	n_{SSR}	n_{morph}
Core *bicostata* (B)						
	3	Hume	3	Tallarook	9	9
		Hume	4	Euroa	12	10
		Hume	5	Eildon	9	10
	4	NE Victoria	6	Mt Granya	8	7
		NE Victoria	7	Shelley	8	7
		NE Victoria	8	Nariel	8	4
		NE Victoria	9	Mt Unicorn	8	5
	5	Canberra	10	Bungongo	13	6
		Canberra	11	Burrinjuck	7	0
Core *bicostata* total					82	58
Core *maidenii* (M)						
	8	Araluen	14	Araluen	16	10
		Araluen	15	Currowan-Monga	23	15
	9	Wadbilliga	16	Wadbilliga North	14	1
		Wadbilliga	17	Belowra Rd	10	7
		Wadbilliga	18	Wadbilliga South	20	7
		Wadbilliga	19	Murrabrine	20	4
		Wadbilliga	20	Kooraban NP	6	0
		Wadbilliga	21	Mt Dromedary (NSW)	13	4
		Wadbilliga	22	Mumbulla SF	20	6
		Wadbilliga	23	Brown Mountain	12	5
	10	SE Forests	24	Tantawangalow	8	1
		SE Forests	25	Big Jack - Rocky Hall	15	4
		SE Forests	26	Yurammie	16	5
		SE Forests	27	Mt Imlay	8	2
Core *maidenii* total					201	71
Core *pseudoglobulus* (P)						
	13	Lakes Entrance	38	Ostlers Rd	9	9
		Lakes Entrance	39	Lake Tyers	14	12
		Lakes Entrance	40	Lakes Entrance	11	11
		Lakes Entrance	41	Metung-Lakes Entrance	6	6
		Lakes Entrance	42	Metung	8	7
Core *pseudoglobulus* total					48	45
Core *globulus* Victoria (G)						
	24	South Gippsland	71	Alberton West-Welshpool-Hedley	14	12
		South Gippsland	72	Toora	8	8
		South Gippsland	73	Fish Creek	6	6
	27	Phillip Island	76	Phillip Island	24	4
Core *globulus* total					52	30
To allocate						
	1	Mt Bryan	1	Mt Bryan	25	19
	2	Mt Cole	2	Mt Cole	19	16
	6	Jenolan	12	Jenolan	13	6
	7	Wollemi	13	Nullo Mountain	20	5
	10	SE Forests	28	north Cann Valley Highway	7	4
		SE Forests	29	Wroxham	8	6
	11	Alfred-Nadgee	30	Nadgee	18	10
		Alfred-Nadgee	31	Maramingo Creek	10	8
		Alfred-Nadgee	32	Alfred NP	9	9
		Alfred-Nadgee	33	mid Cann Valley Highway	6	5
		Alfred-Nadgee	34	south Cann Valley Highway	12	10
	12	Mt Cann	35	Mt Cann	10	8
		Mt Cann	36	Wiebens Hill	8	8

Taxon (Code)	Region		Locality			
	Code	Name	Code	Name	n_{SSR}	n_{morph}
		Mt Cann	37	Murrungowar	8	6
	14	Lerderderg Gorge	43	Lerderderg Gorge	20	16
	15	Buchan	44	Stoney Creek	10	8
		Buchan	45	Cutts Creek Rd	8	7
		Buchan	46	Buchan	11	12
		Buchan	47	Gelantipy	8	8
	16	Omeo	48	Beloka Rd	7	8
		Omeo	49	Omeo	8	8
		Omeo	50	Trapyard Gap	7	6
	17	Mitchell River	51	Cobbannah	10	8
		Mitchell River	52	Peel Gap	10	10
		Mitchell River	53	Heyfield	4	4
	18	Strzelecki Ranges	54	Bowden-Carrajung	2	2
		Strzelecki Ranges	55	Jeeralang	23	25
Grand Total					684	446

Appendix 4. Genetic diversity parameters for *E. globulus*, calculated at the locality level.

Locality (taxon code)	*n*	*A*	R_t	$L_{0.5}$	H_e	H_o	*F*
1 Mt Bryan (B)	24.0	4.6	2.7	0.7	0.57	0.48	0.17
2 Mt Cole (B)	18.7	8.6	3.9	0.1	0.80	0.67	0.17
3 Tallarook (B)	8.8	7.7	4.3	0.0	0.85	0.80	0.06
4 Euroa (B)	11.9	8.2	4.2	0.1	0.84	0.67	0.20
5 Eildon (B)	8.3	6.4	3.9	0.1	0.80	0.72	0.11
6 Mt Granya (B)	7.8	5.4	3.6	0.2	0.76	0.67	0.12
7 Shelley (B)	7.9	5.8	3.7	0.1	0.78	0.60	0.24
8 Nariel (B)	7.7	5.3	3.5	0.2	0.73	0.60	0.19
9 Mt Unicorn (B)	7.9	5.6	3.6	0.3	0.74	0.52	0.32
10 Bungongo (B)	12.3	6.1	3.5	0.1	0.73	0.63	0.13
11 Burrinjuck (B)	6.8	5.6	3.8	0.0	0.79	0.67	0.17
12 Jenolan (B)	12.1	6.2	3.6	0.2	0.76	0.65	0.15
13 Nullo Mountain (B)	19.8	6.8	3.2	0.4	0.68	0.59	0.14
bicostata mean	11.8	6.3	3.7	0.2	0.76	0.64	0.17
14 Araluen (M)	15.8	9.1	4.1	0.1	0.82	0.73	0.11
15 CurrowanMonga (M)	22.8	12.4	4.3	0.0	0.85	0.73	0.14
16 Wadbilliga North (M)	13.7	10.2	4.4	0.1	0.86	0.74	0.14
17 Belowra Rd (M)	10.0	7.7	3.9	0.2	0.78	0.72	0.08
18 Wadbilliga South (M)	19.8	10.9	4.3	0.1	0.84	0.71	0.15
19 Murrabrine (M)	19.7	10.2	4.1	0.1	0.82	0.75	0.08
20 Kooraban NP (M)	5.9	6.9	4.5	0.2	0.85	0.77	0.10
21 MtDromedaryNSW (M)	13.0	9.1	4.2	0.1	0.82	0.79	0.05
22 Mumbulla SF (M)	19.2	10.4	4.2	0.0	0.83	0.69	0.17
23 Brown Mountain (M)	11.7	9.0	4.4	0.0	0.86	0.66	0.24
24 Tantawangalow (M)	8.0	6.4	4.0	0.0	0.82	0.83	-0.02
25 BigJackRockyHall (M)	14.2	9.9	4.4	0.0	0.86	0.72	0.17
26 Yurammie (M)	15.7	10.9	4.5	0.1	0.86	0.76	0.13
27 Mt Imlay (M)	7.8	7.0	4.1	0.2	0.82	0.66	0.22
28 northCann Valley Highway (M)	6.7	7.0	4.4	0.0	0.86	0.75	0.14
29 Wroxham (M)	7.6	6.7	4.2	0.1	0.83	0.71	0.15
maidenii mean	13.2	9.0	4.3	0.1	0.84	0.73	0.13
30 Nadgee (m-p)	17.9	8.1	3.5	0.3	0.72	0.59	0.19
31 Maramingo Creek (m-p)	9.7	8.3	4.4	0.0	0.86	0.72	0.17
32 Alfred NP (m-p)	8.2	7.1	4.1	0.1	0.83	0.72	0.14
33 midCann Valley Highway (m-p)	5.8	5.7	4.0	0.2	0.81	0.76	0.06
34 southCann Valley Highway (m-p)	11.9	9.2	4.3	0.1	0.84	0.74	0.13
m-p intergrade mean	10.7	7.7	4.1	0.2	0.81	0.71	0.14
35 Mt Cann (P)	9.8	8.3	4.3	0.0	0.85	0.70	0.18
36 Wiebens Hill (P)	7.6	7.3	4.3	0.1	0.84	0.73	0.14
37 Murrungowar (P)	7.7	7.0	4.3	0.1	0.84	0.69	0.20
38 Ostlers Rd (P)	8.9	7.3	4.3	0.0	0.85	0.75	0.12
39 Lake Tyers (P)	13.8	9.6	4.4	0.0	0.86	0.75	0.13
40 Lakes Entrance (P)	10.8	8.4	4.4	0.0	0.87	0.80	0.08
41 Metung-Lakes Entrance (P)	5.9	5.0	3.8	0.1	0.80	0.69	0.16

Locality (taxon code)	n	A	R_I	$L_{0.5}$	H_e	H_o	F
42 Metung (P)	7.6	6.2	4.0	0.0	0.82	0.58	0.31
43 Lerderderg Gorge (P)	19.8	9.3	4.0	0.0	0.82	0.68	0.17
pseudoglobulus mean	10.2	7.6	4.2	0.0	0.84	0.71	0.17
44 Stoney Creek (b-m-p)	9.8	8.2	4.3	0.1	0.84	0.71	0.17
45 Cutts Creek Rd (b-m-p)	7.6	5.6	3.6	0.1	0.76	0.69	0.10
46 Buchan (b-m-p)	10.7	9.1	4.5	0.0	0.87	0.67	0.25
47 Gelantipy (b-m-p)	7.7	6.8	4.2	0.0	0.85	0.69	0.20
b-m-p intergrade mean	9.0	7.4	4.2	0.1	0.83	0.69	0.18
48 Beloka Rd (b-p)	7.0	5.8	3.7	0.2	0.76	0.65	0.16
49 Omeo (b-p)	7.9	5.9	3.7	0.2	0.76	0.65	0.15
50 Trapyard Gap (b-p)	7.0	5.3	3.5	0.3	0.72	0.67	0.08
51 Cobbannah (b-p)	9.6	7.7	4.0	0.1	0.79	0.70	0.13
52 Peel Gap (b-p)	9.2	6.4	3.8	0.2	0.78	0.72	0.08
53 Heyfield (b-p)	4.0	4.3	3.7	0.2	0.78	0.61	0.25
b-p intergrade mean	7.5	5.9	3.7	0.2	0.77	0.67	0.14
55 Jeeralang (g)	21.1	8.7	3.9	0.1	0.78	0.69	0.12
56 Lorne (g)	12.9	7.7	4.0	0.2	0.79	0.70	0.11
57 Jamieson Creek (g)	7.7	6.3	3.8	0.2	0.76	0.68	0.12
58 Cape Patton (g)	6.8	5.7	3.8	0.2	0.78	0.72	0.08
59 Parker Spur (g)	7.7	6.6	3.9	0.0	0.81	0.74	0.09
60 Cannan Spur (g)	7.7	6.3	3.7	0.3	0.73	0.67	0.09
61 Otway State Forest (g)	7.4	5.7	3.6	0.3	0.72	0.62	0.14
62 Lavers Hill (g)	7.7	6.1	3.8	0.0	0.80	0.79	0.02
63 North King Island (g)	5.8	4.3	3.1	0.6	0.62	0.46	0.27
64 Central North King Island (g)	4.6	3.9	3.2	0.3	0.69	0.65	0.06
65 Central King Island (g)	19.9	6.8	3.2	0.4	0.66	0.55	0.17
66 South King Island (g)	6.7	4.2	3.0	0.4	0.65	0.64	0.02
67 Little Henty River (g)	7.8	4.8	3.4	0.2	0.73	0.76	-0.05
68 Badgers Creek (g)	7.4	4.8	3.3	0.2	0.70	0.56	0.22
69 Macquarie Harbour (g)	15.8	6.6	3.3	0.3	0.67	0.60	0.11
70 Port Davey (g)	33.8	9.8	3.5	0.2	0.74	0.64	0.13
g intergrade mean	11.3	6.1	3.5	0.3	0.73	0.65	0.11
71 AlbWest_Welshpool_Hedley (G)	13.8	7.4	3.7	0.3	0.76	0.71	0.06
72 Toora (G)	8.0	6.0	3.7	0.3	0.74	0.58	0.22
73 Fish Creek (G)	6.0	5.0	3.6	0.3	0.75	0.72	0.04
74 Tidal River (G)	29.7	7.3	3.3	0.4	0.69	0.76	-0.10
75 Wilson's Prom Lighthouse (G)	28.6	5.8	3.0	0.3	0.65	0.64	0.00
76 Phillip Island (G)	23.3	10.2	3.8	0.2	0.77	0.71	0.08
77 RechercheBay (G)	25.4	9.2	3.9	0.1	0.79	0.68	0.15
78 SthBrunyIsland (G)	7.9	5.2	3.6	0.1	0.77	0.68	0.12
79 Dover (G)	6.7	6.2	4.0	0.1	0.79	0.73	0.09
80 Geeveston (G)	9.3	6.3	3.8	0.0	0.80	0.79	0.02
81 Tinderbox (G)	8.7	6.0	3.7	0.2	0.77	0.65	0.16
83 Maria Island South (G)	28.6	10.4	4.0	0.1	0.81	0.69	0.15
84 Maria Island North (G)	5.9	4.9	3.6	0.1	0.78	0.70	0.11
85 Triabunna (G)	7.4	6.6	4.0	0.0	0.82	0.71	0.14

Locality (taxon code)	n	A	R_t	$L_{0.5}$	H_e	H_o	F
86 MtDromedaryTas (G)	18.0	8.4	3.8	0.1	0.79	0.73	0.08
87 Platform Peak (G)	16.3	8.0	3.8	0.1	0.79	0.71	0.10
88 Jericho (G)	5.4	4.3	3.3	0.2	0.74	0.65	0.12
89 Mayfield South (G)	5.2	4.3	3.3	0.4	0.70	0.61	0.15
90 Mayfield North (G)	7.1	5.3	3.6	0.3	0.74	0.58	0.23
91 Cape Tourville Dwarf (G)	19.1	8.3	3.8	0.1	0.79	0.68	0.14
92 Cape Tourville Tall (G)	35.3	11.9	4.1	0.0	0.81	0.72	0.12
93 Pepper Hill (G)	7.6	5.2	3.4	0.3	0.70	0.64	0.09
94 Humbug Hill (G)	8.8	5.8	3.6	0.2	0.76	0.71	0.07
95 Cape Barren Island (G)	25.9	9.4	3.8	0.1	0.78	0.67	0.15
97 Flinders Island central (G)	14.6	7.9	3.9	0.0	0.80	0.72	0.10
globulus mean	14.9	7.0	3.7	0.2	0.76	0.69	0.10
E. globulus mean	12.2	7.1	3.8	0.2	0.78	0.68	0.13

n = mean number of individual trees per locus per population, A = mean number of alleles per locus, R_t = allelic richness, $L_{0.5}$= proportion of loci with an allele > 0.5 in frequency H_e = expected heterozygosity, H_o = observed heterozygosity, F = Wright's Fixation Index. For taxon codes see Table 1.2.

REFERENCES

Adam H, Jouannic S, Morcillo F, Richaud F, Duval Y, Tregear JW (2006) MADS box genes in oil palm (*Elaeis guineensis*): Patterns in the evolution of the *SQUAMOSA*, *DEFICIENS*, *GLOBOSA*, *AGAMOUS*, and *SEPALLATA* subfamilies. *Journal of Molecular Evolution* **62**, 15-31.

Albrecht C, Geurts R, Lapeyrie F, Bisseling T (1998) Endomycorrhizae and rhizobial Nod factors both require SYM8 to induce the expression of the early nodulin genes *PsENOD5* and *PsENOD12A*. *Plant Journal* **15**, 605-614.

Aldrich PR, Parker GR, Michler CH, Romero-Severson J (2003) Whole-tree silvic identifications and the microsatellite genetic structure of a red oak species complex in an Indiana old-growth forest. *Canadian Journal of Forest Research* **33**, 2228-2237.

Alwee SS, Van der Linden CG, Van der Schoot J, de Folter S, Angenent C, Cheah SC, Smulders MJM (2006) Characterization of oil palm MADS box genes in relation to the mantled flower abnormality. *Plant Cell Tissue and Organ Culture* **85**, 331-344.

Apiolaza LA, Potts BM, Gore PL (2001) Genetic control of peak flowering time of *Eucalyptus globulus*. In 'IUFRO International Symposium on Developing the Eucalypt of the Future'. Valdivia, Chile. (Ed. S Barros) p. 155. (INFOR)

Aradhya KM, Muellerdombois D, Ranker TA (1991) Genetic evidence for recent and incipient speciation in the evolution of Hawaiian *Metrosideros* (Myrtaceae). *Heredity* **67**, 129-138.

Ashton DH (1975) Studies of the flowering behaviour in *Eucalyptus regnans* F. Muell. *Australian Journal of Botany* **23**, 399-411.

Barbour RC, Baker SC, O'Reilly-Wapstra JM, Harvest TMA, Potts BM (2008a) A footprint of tree-genetics on the biota of the forest floor. *Oikos*, in press.

Barbour RC, Forster LG, Baker SC, Potts BM (2009a) Biodiversity consequences of genetic variation in bark characteristics within a foundation tree species. *Conservation Biology*, in press.

Barbour RC, O'Reilly-Wapstra JM, de Little DW, Jordan GJ, Steane DA, Humphreys JR, Bailey JK, Whitham TG, Potts BM (2009b) A geographic mosaic of genetic variation within a foundation tree species and its community-level consequence. *Ecology*, in press.

Barbour RC, Otahal Y, Vaillancourt RE, Potts BM (2008b) Assessing the risk of pollen-mediated gene flow from exotic *Eucalyptus globulus* plantations into native eucalypt populations of Australia. *Biological Conservation* **141**, 896-907.

Barbour RC, Potts BM, Vaillancourt RE, Tibbits WN (2006) Gene flow between introduced and native *Eucalyptus* species: flowering asynchrony as a barrier to F1 hybridisation between exotic *E. nitens* and native Tasmanian *Symphyomyrtus* species. *Forest Ecology and Management* **226**, 9-21.

Barrows TT, Stone JO, Fifield LK, Cresswell RG (2001) Late Pleistocene glaciation of the Kosciuszko Massif, Snowy Mountains, Australia. *Quaternary Research* **55**, 179-189.

Battey NH (2000) Aspects of seasonality. *Journal of Experimental Botany* **51**, 1769-1780.

Beckstrom-Sternberg SM, Rieseberg LH, Doan K (1991) Gene lineage analysis in populations of *Helianthus niveus* and *H. petiolaris* (Asteraceae). *Plant Systematics and Evolution* **175**, 125-138.

Bicer BT, Sakar D (2008) Heritability and path analysis of some economical characteristics in lentil. *Journal of Central European Agriculture* **9**, 175-179.

Birky CW (1995) Uniparental inheritance of mitochondrial and chloroplast genes - mechanisms and evolution. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 11331-11338.

Birtchnell MJ, Gibson M (2006) Long-term flowering patterns of melliferous *Eucalyptus* (Myrtaceae) species. *Australian Journal of Botany* **54**, 745-754.

Bloomfield J (2008) Molecular variation between Tasmanian *Eucalyptus obliqua* populations. Honours thesis, University of Tasmania.

Bohlenius H, Huang T, Charbonnel-Campaa L, Brunner AM, Jansson S, Strauss SH, Nilsson O (2006) CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees. *Science* **312**, 1040-1043.

Boland DJ, Brooker MIH, Turnbull JW, Kleinig DA (1980) '*Eucalyptus* seed.' (CSIRO: Melbourne)

Boss PK, Bastow RM, Mylne JS, Dean C (2004) Multiple pathways in the decision to flower: Enabling, promoting, and resetting. *Plant Cell* **16**, S18-S31.

Boyer WD (1978) Heat accumulation: an easy way to anticipate the flowering of southern pines. *Journal of Forestry* **76**, 20-23.

Bradstock RA, Gill AM, Hastings SM, Moore PHR (1994) Survival of serotinous seedbanks during bushfires: comparative studies of *Hakea* species from southern Australia. *Australian Journal of Ecology* **19**, 276-282.

Brill EM, Watson JM (2004) Ectopic expression of a *Eucalyptus grandis* SVP orthologue alters the flowering time of *Arabidopsis thaliana*. *Functional Plant Biology* **31**, 217-224.

Brondani RPV, Brondani C, Grattapaglia D (2002) Towards a genus-wide reference linkage map for *Eucalyptus* based exclusively on highly informative microsatellite markers. *Molecular Genetics and Genomics* **267**, 338-347.

Brondani RPV, Brondani C, Tarchini R, Grattapaglia D (1998) Development, characterization and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophylla*. *Theoretical and Applied Genetics* **97**, 816-827.

Brooker MIH (2000) A new classification of the genus *Eucalyptus* L'Her. (Myrtaceae). *Australian Systematic Botany* **13**, 79-148.

Brooker MIH, Kleinig DA (1999) 'Field Guide to Eucalypts. Volume 1. South-eastern Australia 2nd Ed.' (Bloomings Books: Hawthorn, Victoria)

Brunner AM, Nilsson O (2004) Revisiting tree maturation and floral initiation in the poplar functional genomics era. *New Phytologist* **164**, 43-51.

Brunner AM, Yakovlev IA, Strauss SH (2004) Validating internal controls for quantitative plant gene expression studies. *BMC Plant Biology* **4**, 14.

Bustin SA (2000) Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *Journal of Molecular Endocrinology* **25**, 169-193.

Butcher PA, McDonald MW, Bell JC (2009) Congruence between environmental parameters, morphology and genetic structure in Australia's most widely distributed eucalypt, *Eucalyptus camaldulensis*. *Tree Genetics & Genomes* **5**, 189-210.

Butcher PA, Skinner AK, Gardiner CA (2005) Increased inbreeding and inter-species gene flow in remnant populations of the rare *Eucalyptus benthamii*. *Conservation Genetics* **6**, 213-226.

Byrne M (1999) High genetic identities between three oil mallee taxa, *Eucalyptus kochii* ssp. *kochii* ssp. *plenissima* and *E. horistes*, based on nuclear RFLP analysis. *Heredity* **82**, 205-211.

Byrne M (2008) Phylogeny, diversity and evolution of eucalypts. In 'Plant Genome: Biodiversity and Evolution. Volume 1, Part E: Phanerogams - Angiosperm'. (Eds AK Sharma and A Sharma) pp. 303-346. (Science Publishers: Enfield)

Byrne M, Elliott CP, Yates CJ, Coates DJ (2008) Maintenance of high pollen dispersal in *Eucalyptus wandoo*, a dominant tree of the fragmented agricultural region in Western Australia. *Conservation Genetics* **9**, 97-105.

Byrne M, MacDonald B, Coates D (2002) Phylogeographical patterns in chloroplast DNA variation within the *Acacia acuminata* (Leguminosae : Mimosoideae) complex in Western Australia. *Journal of Evolutionary Biology* **15**, 576-587.

Byrne M, Moran GF, Tibbits WN (1993) Restriction map and maternal inheritance of chloroplast DNA in *Eucalyptus nitens*. *Journal of Heredity* **84**, 218-220.

Caetano S, Nusboumer L, Naciri Y (2008) Chloroplast and microsatellite markers in *Astronium urundeuva* (Allemao) Engl. and close species of Anacardiaceae: toward the definition of a species complex? *Candollea* **63**, 115-130.

Carlsbecker A, Tandre K, Johanson U, Englund M, Engstrom P (2004) The MADS-box gene *DAL1* is a potential mediator of the juvenile-to-adult transition in Norway spruce (*Picea abies*). *Plant Journal* **40**, 546-557.

Chambers PGS, Potts BM, Tilyard PG (1997) The genetic control of flowering precocity in *Eucalyptus globulus*. *Silvae Genetica* **46**, 207-214.

Chao CT, Parfitt DE (2003) Genetic analyses of phenological traits of pistachio (*Pistacia vera* L.). *Euphytica* **129**, 345-349.

Chapuis M-P, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. *Mol Biol Evol* **24**, 621-631.

Cheng CH, Seal AG, Murphy SJ, Lowe RG (2006) Variability and inheritance of flowering time and duration in Actinidia chinensis (kiwifruit). *Euphytica* **147**, 395-402.

Cho SK, Larue CT, Chevalier D, Wang HC, Jinn TL, Zhang SQ, Walker JC (2008) Regulation of floral organ abscission in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 15629-15634.

Christensen JH, Hewitson B, Busuioc A, Chen A, Gao X, Held I, Jones R, Kolli RK, Kwon W-T, Laprise R, Rueda VM, Mearns L, Menéndez CG, Räisänen J, Rinke A, Sarr A, Whetton P (2007) Regional Climate Projections. In 'Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change'. (Eds S. Solomon, D Qin, M Manning, Z Chen, M Marquis, KB Averyt, M Tignor and HL Miller). (Cambridge University Press: Cambridge, United Kingdom and New York, NY, USA)

Chybicki IJ, Burczyk J (2009) Simultaneous estimation of null alleles and inbreeding coefficients. *Journal of Heredity* **100**, 106-113.

Clifford HT, Drake WE (1981) Pollination and dispersal in eastern Australian heathlands. In 'Ecosystems of the World 9B. Heathlands and Related Shrublands. Analytical Studies'. (Ed. RL Specht) pp. 39-49. (Elsevier: Amsterdam)

Cook LG, Morris DC, Edwards RD, Crisp MD (2008) Reticulate evolution in the natural range of the invasive wetland tree species *Melaleuca quinquenervia*. *Molecular Phylogenetics and Evolution* **47**, 506-522.

Corbesier L, Vincent C, Jang SH, Fornara F, Fan QZ, Searle I, Giakountis A, Farrona S, Gissot L, Turnbull C, Coupland G (2007) FT protein movement contributes to

long-distance signaling in floral induction of *Arabidopsis*. *Science* **316**, 1030-1033.

Crawford NG (2009) SMOGD: Software for the Measurement of Genetic Diversity version 1.2.0. *Molecular Ecology Resources.* **In Prep.**

Dakin EE, Avise JC (2004) Microsatellite null alleles in parentage analysis. *Heredity* **93**, 504-509.

DeVries DP (1967) Phenological stages in sweet cherry with regard to pre-selection. *Euphytica* **16**, 177-182.

Diaz EC, Martin F, Tagu D (1996) Eucalypt alpha-tubulin: cDNA cloning and increased level of transcripts in ectomycorrhizal root system. *Plant Molecular Biology* **31**, 905-910.

Dicenta F, Garcia JE, Carbonell EA (1993) Heritability of flowering, productivity and maturity in almond. *Journal of Horticultural Science* **68**, 113-120.

Dobzhansky T (1970) 'Genetics of the Evolutionary Process.' (Columbia University Press: New York)

Dornelas MC, Rodriguez APM (2005a) Identifying *Eucalyptus* expressed sequence tags related to *Arabidopsis* flowering-time pathway genes. *Brazilian Journal of Plant Physiology* **17**, 255-266.

Dornelas MC, Rodriguez APM (2005b) The rubber tree (*Hevea brasiliensis* Muell. Arg.) homologue of the *LEAFY/FLORICAULA* gene is preferentially expressed in both male and female floral meristems. *Journal of Experimental Botany* **56**, 1965-1974.

Dornelas MC, Rodriguez APM (2006) The tropical cedar tree (*Cedrela fissilis* Vell., Meliaceae) homolog of the *Arabidopsis LEAFY* gene is expressed in reproductive tissues and can complement *Arabidopsis leafy* mutants. *Planta* **223**, 306-314.

Doyle JJ, Doyle JL (1990) Extraction of plant DNA from fresh tissue. *Focus* **12**, 13-15.

Dressler RL, Dodson CH (1960) Classification and phylogeny in the Orchidaceae. *Annals Missouri Botanical Gardens.* **47**, 25-68.

Dugardeyn J, Van der Straeten D (2008) Ethylene: Fine-tuning plant growth and development by stimulation and inhibition of elongation. *Plant Science* **175**, 59-70.

Duminil J, Caron H, Scotti I, Cazal SO, Petit RJ (2006) Blind population genetics survey of tropical rainforest trees. *Molecular Ecology* **15**, 3505-3513.

Dungey HS, Potts BM, Whitham TG, Li HF (2000) Plant genetics affects arthropod community richness and composition: Evidence from a synthetic eucalypt hybrid population. *Evolution* **54**, 1938-1946.

Dutkowski GW, Potts BM (1999) Geographic patterns of genetic variation in *Eucalyptus globulus* ssp. *globulus* and a revised racial classification. *Australian Journal of Botany* **47**, 237-263.

Eckert CG, Samis KE, Lougheed SC (2008) Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. *Molecular Ecology* **17**, 1170-1188.

El Mousadik A, Petit RJ (1996) High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L) Skeels] endemic to Morocco. *Theoretical and Applied Genetics* **92**, 832-839.

Eldridge K, Davidson J, Harwood C, van Wyk G (1993) 'Eucalypt domestication and breeding.' (Clarendon Press: Oxford)

Elliott CP, Byrne M (2004) Phylogenetics and the conservation of rare taxa in the *Eucalyptus angustissima* complex in Western Australia. *Conservation Genetics* **5**, 39-47.

Elo A, Lemmetyinen J, Novak A, Keinonen K, Porali I, Hassinen M, Sopanen T (2007) *BpMADS4* has a central role in inflorescence initiation in silver birch (*Betula pendula*). *Physiologia Plantarum* **131**, 149-158.

Elo A, Lemmetyinen J, Turunen ML, Tikka L, Sopanen T (2001) Three MADS-box genes similar to *APETALA1* and *FRUITFULL* from silver birch (*Betula pendula*). *Physiologia Plantarum* **112**, 95-103.

Endler J (1977) 'Geographic Variation, Speciation and Clines.' (Princeton University Press: Princeton, New Jersey)

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**, 2611-2620.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes - application to human mitochondrial-DNA restriction data. *Genetics* **131**, 479-491.

Falconer DS, MacKay TFC (1996) 'Introduction to Quantitative Genetics.' (Longman Scientific and Technical: New York)

Fernandez DE, Heck GR, Perry SE, Patterson SE, Bleecker AB, Fang SC (2000) The embryo MADS domain factor *AGL15* acts postembryonically: Inhibition of perianth senescence and abscission via constitutive expression. *Plant Cell* **12**, 183-197.

Fernando DD, Zhang SL (2006) Constitutive expression of the *SAP1* gene from willow (*Salix discolor*) causes early flowering in *Arabidopsis thaliana*. *Development Genes and Evolution* **216**, 19-28.

Ferrante A, Mensuali-Sodi A, Serra G, Tognoni F (2002) Effects of ethylene and cytokinins on vase life of cut *Eucalyptus parvifolia* Cambage branches. *Plant Growth Regulation* **38**, 119-125.

Foo E, Buillier E, Goussot M, Foucher F, Rameau C, Beveridge CA (2005) The branching gene *RAMOSUS1* mediates interactions among two novel signals 464 and auxin in pea. *Plant Cell* **17**, 464-474.

Foster SA, McKinnon GE, Steane DA, Potts BM, Vaillancourt RE (2007) Parallel evolution of dwarf ecotypes in the forest tree *Eucalyptus globulus*. *New Phytologist* **175**, 370-380.

Foucher F, Morin J, Courtiade J, Cadioux S, Ellis N, Banfield MJ, Rameau C (2003) *DETERMINATE* and *LATE FLOWERING* are two *TERMINAL FLOWER1/CENTRORADIALIS* homologs that control two distinct phases of flowering initiation and development in pea. *Plant Cell* **15**, 2742-2754.

Frankel O, Soule M (1981) 'Conservation and Evolution.' (Cambridge University Press: Cambridge)

Freeman JS, Jackson HD, Steane DA, McKinnon GE, Dutkowski GW, Potts BM, Vaillancourt RE (2001) Chloroplast DNA phylogeography of *Eucalyptus globulus*. *Australian Journal of Botany* **49**, 585-596.

Galouchko V (2007) '3DField.' (Available online at http://field.hypermart.net/).

Garcia-Mozo H, Galan C, Jato V, Belmonte J, de la Guardia CD, Fernandez D, Gutierrez M, Aira MJ, Roure JM, Ruiz L, Trigo MM, Dominguez-Vilches E (2006) *Quercus* pollen season dynamics in the Iberian Peninsula: Response to meteorological parameters and possible consequences of climate change. *Annals of Agricultural and Environmental Medicine* **13**, 209-224.

Gardiner C, Crawford D (1987) '1987 Seed Collections of *Eucalyptus globulus* subsp. *globulus* for tree improvement purposes. Unpublished report.' Tree Seed Centre, CSIRO Division of Forest Research, Canberra.

Gardiner C, Crawford D (1988) '1988 Seed Collections of *Eucalyptus globulus* subsp. *globulus* for tree improvement purposes. Unpublished report.' Tree Seed Centre, CSIRO Division of Forest Research, Canberra.

Gaston KJ, Fuller RA (2008) Commonness, population depletion and conservation biology. *Trends in Ecology & Evolution* **23**, 14-19.

Gilmour AR, Thompson R, Cullis BR (1995) Average Information REML, an efficient algorithm for variance parameter estimation in linear mixed models. *Biometrics* **51**, 1440-1450.

Gomory D, Paule L, Vysny J (2007) Patterns of allozyme variation in western Eurasian *Fagus*. *Botanical Journal of the Linnean Society* **154**, 165-174.

Gomory D, Yakovlev I, Zhelev P, Jedinakova J, Paule L (2001) Genetic differentiation of oak populations within the *Quercus robur/Quercus petraea* complex in Central and Eastern Europe. *Heredity* **86**, 557-563.

Gonzalez-Parrado Z, Fuertes-Rodriguez CR, Vega-Maray AM, Valencia-Barrera RM, Rodriguez-Rajo FJ, Fernandez-Gonzalez D (2006) Chilling and heat requirements for the prediction of the beginning of the pollen season of *Alnus glutinosa* (L.) Gaertner in Ponferrada (Leon, Spain). *Aerobiologia* **22**, 47-53.

Gore PL, Potts BM (1995) The genetic control of flowering time in *Eucalyptus globulus*, *E. nitens* and their F1 hybrids. In 'Eucalypt plantations: improving fibre yield and quality, Proc. CRCTHF-IUFRO Conf., Hobart, 19-24 Feb.' (Eds BM Potts, NMG Borralho, JB Reid, RN Cromer, WN Tibbits and CA Raymond) pp. 241-242. (CRC for Temperate Hardwood Forestry, Hobart, Tasmania)

Goudet J (2002) 'FSTAT, a program to estimate and test gene diversities and fixation indices.' (Available online at http://www2.unil.ch/popgen/softwares/fstat.htm).

Grattapaglia D, Kirst M (2008) *Eucalyptus* applied genomics: from gene sequences to breeding tools. *New Phytologist* **179**, 911-929.

Griffin AR (1980) Floral phenology of a stand of mountain ash (*Eucalyptus regnans* F. Muell.) in Gippsland, Victoria. *Australian Journal of Botany* **28**, 393–404.

Griffin AR (1982) Pollination ecology of eucalypts—a framework for study. In 'Pollination '82'. (Eds EG Williams, RB Knox, JH Gilbert and P Bernhardt.) pp. 42–56. (University of Melbourne: Plant Cell Biology Research Centre, School of Botany.)

Griffin AR, Burgess IP, Wolf L (1988) Patterns of natural and manipulated hybridization in the genus *Eucalyptus* L'Herit - a review. *Australian Journal of Botany* **36**, 41-66.

Griffin AR, Whiteman P, Rudge T, Burgess IP, Moncur M (1993) Effect of paclobutrazol on flower bud production and vegetative growth in two species of *Eucalyptus*. *Canadian Journal of Forest Research* **23**, 640-647.

Guillaumot D, Lelu-Walter MA, Germot A, Meytraud F, Gastinel L, Riou-Khamlichi C (2008) Expression patterns of *LmAP2L1* and *LmAP2L2* encoding two-APETALA2 domain proteins during somatic embryogenesis and germination of hybrid larch (*Larix* x marschlinsii). *Journal of Plant Physiology* **165**, 1003-1010.

Gyllenstrand N, Clapham D, Kallman T, Lagercrantz U (2007) A Norway Spruce *FLOWERING LOCUS T* homolog is implicated in control of growth rhythm in conifers. *Plant Physiology* **144**, 248.

Hartl DL, Clark AG (1997) 'Principles of Population Genetics.' (Sinauer Associates Inc.: Sunderland, Massachusetts)

Hasan O, Reid JB (1995) Reduction of generation time in *Eucalyptus globulus*. *Plant Growth Regulation* **17**, 53-60.

Hattasch C, Flachowsky H, Kapturska D, Hanke MV (2008) Isolation of flowering genes and seasonal changes in their transcript levels related to flower induction and initiation in apple (*Malus domestica*). *Tree Physiology* **28**, 1459-1466.

Hayama R, Yokoi S, Tamaki S, Yano M, Shimamoto K (2003) Adaptation of photoperiodic control pathways produces short-day flowering in rice. *Nature* **422**, 719-722.

Hecht V, Foucher F, Ferrandiz C, Macknight R, Navarro C, Morin J, Vardy ME, Ellis N, Beltran JP, Rameau C, Weller JL (2005) Conservation of *Arabidopsis* flowering genes in model legumes. *Plant Physiology* **137**, 1420-1434.

Hedrick PW (1999) Perspective: Highly variable loci and their interpretation in evolution and conservation. *Evolution* **53**, 313-318.

Hedrick PW (2005) A standardized genetic differentiation measure. *Evolution* **59**, 1633-1638.

Heller R, Siegismund HR (2009) Relationship between three measures of genetic differentiation G_{ST}, D_{est} and G'_{ST}: how wrong have we been? *Molecular Ecology* **18**, 2080-2083.

Hess WJ, Stoynoff NA (1998) Taxonomic status of *Quercus acerifolia* (Fagaceae) and a morphological comparison of four members of the *Quercus shumardii* complex. *Systematic Botany* **23**, 89-100.

Hetherington S, Jones KM, Koen TB (1992) Stimulation of bud production in *Eucalyptus globulus* by paclobutrazol application. *South African Forestry Journal* **160**.

Hines B, Byrne M (2001) Genetic differentiation between mallee and tree forms in the *Eucalyptus loxophleba* complex. *Heredity* **87**, 566-572.

Hingston AB, Gartrell BD, Pinchbeck G (2004a) How specialized is the plant-pollinator association between *Eucalyptus globulus* ssp. *globulus* and the swift parrot *Lathamus discolor*? *Austral Ecology* **29**, 624-630.

Hingston AB, Potts BM, McQuillan PB (2004b) The swift parrot, *Lathamus discolor* (Psittacidae), social bees (Apidae) and native insects as pollinators of *Eucalyptus globulus* ssp. *globulus* (Myrtaceae). *Australian Journal of Botany* **52**, 371-379.

Hofer J, Turner L, Hellens R, Ambrose M, Matthews P, Michael A, Ellis N (1997) *UNIFOLIATA* regulates leaf and flower morphogenesis in pea. *Current Biology* **7**, 581-587.

Hokanson SC, Isebrands JG, Jensen RJ, Hancock JF (1993) Isozyme variation in oaks of the Apostle Islands in Wisconsin - Genetic structure and levels of inbreeding in *Quercus rubra* and *Q. ellipsoidalis* (Fagaceae). *American Journal of Botany* **80**, 1349-1357.

Holman JE, Hughes JM, Fensham RJ (2003) A morphological cline in *Eucalyptus*: a genetic perspective. *Molecular Ecology* **12**, 3013-3025.

Hope GS (1994) Quarternary vegetation. In 'History of the Australian vegetation: Cretaceous to recent'. (Ed. RS Hill) pp. 368-389. (Cambridge University Press: Cambridge)

House SM (1997) Reproductive biology of eucalypts. In 'Eucalypt Ecology: Individuals to Ecosystems'. (Eds J Williams and J Woinarski) pp. 30-55. (Cambridge University Press: Cambridge)

Hsu CY, Liu YX, Luthe DS, Yuceer C (2006) Poplar *FT2* shortens the juvenile phase and promotes seasonal flowering. *Plant Cell* **18**, 1846-1861.

Igasaki T, Watanabe Y, Nishiguchi M, Kotoda N (2008) The *FLOWERING LOCUS T/TERMINAL FLOWER 1* family in Lombardy poplar. *Plant and Cell Physiology* **49**, 291-300.

Ingvarsson PK (unpubl.) 'Multilocus patterns of nucleotide polymorphism and the demographic history of *Populus tremula*.' Sequences submitted to GenBank: Accessions EU824410-EU824433.

Izawa T, Takahashi Y, Yano M (2003) Comparative biology comes into bloom: genomic and genetic comparison of flowering pathways in rice and *Arabidopsis*. *Current Opinion in Plant Biology* **6**, 113-120.

Jackson HD, Steane DA, Potts BM, Vaillancourt RE (1999) Chloroplast DNA evidence for reticulate evolution in *Eucalyptus* (Myrtaceae). *Molecular Ecology* **8**, 739-751.

Jackson WD (1999) Vegetation types. In 'Vegetation of Tasmania'. (Eds JB Reid, RS Hill, MJ Brown and MJ Hovenden) pp. 89–124. (Australian Biological Resource Study: Melbourne)

Jager M, Hassanin A, Manuel M, Le Guyader H, Deutsch J (2003) MADS-box genes in *Ginkgo biloba* and the evolution of the *AGAMOUS* family. *Molecular Biology and Evolution* **20**, 842-854.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**, 1801-1806.

Jarvinen P, Lemmetyinen J, Savolainen O, Sopanen T (2003) DNA sequence variation in *BpMADS2* gene in two populations of *Betula pendula*. *Molecular Ecology* **12**, 369-384.

Johnson LAS (1976) Problems of species and genera in *Eucalyptus* (Myrtaceae). *Plant Systematics and Evolution* **125**, 155-167.

Johnson PR, Ecker JR (1998) The ethylene gas signal transduction pathway: A molecular perspective. *Annual Review of Genetics* **32**, 227-254.

Johnson X, Brcich T, Dun EA, Goussot M, Haurogne K, Beveridge CA, Rameau C (2006) Branching genes are conserved across species. Genes controlling a novel signal in pea are coregulated by other long-distance signals. *Plant Physiology* **142**, 1014-1026.

Jones RC, McKinnon GE, Potts BM, Vaillancourt RE (2005) Genetic diversity and mating system of an endangered tree *Eucalyptus morrisbyi*. *Australian Journal of Botany* **53**, 367-377.

Jones RC, Steane DA, Potts BM, Vaillancourt RE (2002) Microsatellite and morphological analysis of *Eucalyptus globulus* populations. *Canadian Journal of Forest Research* **32**, 59-66.

Jones TH, Steane DA, Jones RC, Pilbeam D, Vaillancourt RE, Potts BM (2006) Effects of domestication on genetic diversity in *Eucalyptus globulus*. *Forest Ecology and Management* **234**, 78-84.

Jones TH, Vaillancourt RE, Potts BM (2007) Detection and visualization of spatial genetic structure in continuous *Eucalyptus globulus* forest. *Molecular Ecology* **16**, 697-707.

Jordan GJ, Borralho NMG, Tilyard P, Potts BM (1994) Identification of races in *Eucalyptus globulus* ssp. *globulus* based on growth traits in Tasmania and geographic distribution. *Silvae Genetica* **43**, 292-298.

Jordan GJ, Potts BM, Chalmers P, Wiltshire RJE (2000) Quantitative genetic evidence that the timing of vegetative phase change in *Eucalyptus globulus* ssp. *globulus* is an adaptive trait. *Australian Journal of Botany* **48**, 561-567.

Jordan GJ, Potts BM, Kirkpatrick JB, Gardiner C (1993) Variation in the *Eucalyptus globulus* complex revisited. *Australian Journal of Botany* **41**, 763-785.

Jordan GJ, Potts BM, Wiltshire RJE (1999) Strong, independent, quantitative genetic control of the timing of vegetative phase change and first flowering in *Eucalyptus globulus* ssp. *globulus* (Tasmanian Blue Gum). *Heredity* **83**, 179-187.

Jost L (2008) G_{ST} and its relatives do not measure differentiation. *Molecular Ecology* **17**, 4015-4026.

Kashani N, Dodd RS (2002) Genetic differentiation of two California red oak species, *Quercus parvula* var. *shreveii* and *Q. wislizeni*, based on AFLP genetic markers. General Technical Report - Pacific Southwest Research Station, USDA Forest Service, 417-426.

Keatley MR, Hudson IL (1998) The influence of fruit and bud volumes on eucalypt flowering - an exploratory analysis. *Australian Journal of Botany* **46**, 281-304.

Kennington WJ, James SH (1998) Allozyme and morphometric variation in two closely related mallee species from Western Australia, *Eucalyptus argutifolia* and *E. obtusiflora* (Myrtaceae). *Australian Journal of Botany* **46**, 173-186.

Kim S, Koh J, Ma H, Hu Y, Endress PK, Hauser BA, Buzgo M, Soltis PS, Soltis DE (2005a) Sequence and expression studies of A-, B-, and E-class MADS-box homologues in *Eupomatia* (Eupomatiaceae): Support for the bracteate origin of the calyptra. *International Journal of Plant Sciences* **166**, 185-198.

Kim S, Koh J, Yoo MJ, Kong HZ, Hu Y, Ma H, Soltis PS, Soltis DE (2005b) Expression of floral MADS-box genes in basal angiosperms: implications for the evolution of floral regulators. *Plant Journal* **43**, 724-744.

Kirkpatrick J (1975) 'Geographical variation in *Eucalyptus globulus*.' Forestry and Timber Bureau, Australia.

Kirkpatrick JB (1973) Geographic variation in *Eucalyptus globulus* Labill. including studies on the distribution and taxonomy of the species. PhD thesis, University of Melbourne.

Kirkpatrick JB (1974) The numerical intraspecific taxonomy of *Eucalyptus globulus* Labill. (Myrtaceae). *Botanical Journal of the Linnean Society* **69**, 89-104.

Kirkpatrick JB, Fowler M (1998) Locating likely glacial forest refugia in Tasmania using palynological and ecological information to test alternative climatic models. *Biological Conservation* **85**, 171-182.

Kirkpatrick JB, Simmons D, Parsons RF (1973) The relationship of some populations involving *Eucalyptus cypellocarpa* and *E. globulus* to the problem of phantom hybrids. *New Phytologist* **72**, 867-876.

Kobayashi Y, Kaya H, Goto K, Iwabuchi M, Araki T (1999) A pair of related genes with antagonistic roles in mediating flowering signals. *Science* **286**, 1960-1962.

Kojima S, Takahashi Y, Kobayashi Y, Monna L, Sasaki T, Araki T, Yano M (2002) *Hd3a*, a rice ortholog of the *Arabidopsis FT* gene, promotes transition to flowering downstream of *Hd1* under short-day conditions. *Plant and Cell Physiology* **43**, 1096-1105.

Kotoda N, Iwanami H, Takahashi S, Abe K (2006) Antisense expression of *MdTFL1*, a *TFL1*-like gene, reduces the juvenile phase in apple. *Journal of the American Society for Horticultural Science* **131**, 74-81.

Kotoda N, Wada M (2005) *MdTFL1*, a *TFL1*-like gene of apple, retards the transition from the vegetative to reproductive phase in transgenic *Arabidopsis*. *Plant Science* **168**, 95-104.

Kyozuka J, Harcourt R, Peacock WJ, Dennis ES (1997) *Eucalyptus* has functional equivalents of the *Arabidopsis AP1* gene. *Plant Molecular Biology* **35**, 573-584.

Ladiges PY (1997) Phylogenetic history and classification of eucalypts. In 'Eucalypt Ecology: Individuals to Ecosystems'. (Eds J Williams and J Woinarski) pp. 16-29. (Cambridge University Press: Cambridge)

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948.

Lefevre F (2004) Human impacts on forest genetic resources in the temperate zone: an updated review. *Forest Ecology and Management* **197**, 257-271.

Lemmetyinen J, Hassinen M, Elo A, Porali I, Keinonen K, Makela H, Sopanen T (2004) Functional characterization of *SEPALLATA3* and *AGAMOUS* orthologues in silver birch. *Physiologia Plantarum* **121**, 149-162.

Lemmetyinen J, Pennanen T, Lannenpaa M, Sopanen T (2001) Prevention of flower formation in dicotyledons. *Molecular Breeding* **7**, 341-350.

Levin DA (1978) The origin of isolating mechanisms in flowering plants. *Evolutionary Biology* **11**, 185-317.

Lewis PO, Zaykin D (2001) 'Genetic Data Analysis: computer program for the analysis of allelic data.' (Available online at http://hydrodictyon.eeb.uconn.edu/people/plewis/software.php).

Li JZ, Absher DM, Tang H, Southwick AM, Casto AM, Ramachandran S, Cann HM, Barsh GS, Feldman M, Cavalli-Sforza LL, Myers RM (2008) Worldwide human relationships inferred from genome-wide patterns of variation. *Science* **319**, 1100-1104.

MacDonald AC, Borralho NMG, Potts BM (1997) Genetic variation for growth and wood density in *Eucalyptus globulus* ssp. *globulus* in Tasmania (Australia). *Silvae Genetica* **46**, 236-241.

Machado CD, Teixeira NJP, Freire FR, Rocha MD, Gomes RLF (2008) Identification of cowpea genotypes for earliness, plant architecture and grain yield. *Revista Ciencia Agronomica* **39**, 114-123.

Magee JW, Miller GH (1998) Lake Eyre palaeohydrology from 60 ka to the present: beach ridges and glacial maximum aridity. *Palaeogeography Palaeoclimatology Palaeoecology* **144**, 307-329.

Mahna N, Dreesen R, Baghban Kohneh Rouz B, Ghareyazie B, Valizadeh M, Grigorian V, Keulemans J (2006) Isolation and characterization of a MADS-box type gene from apple (*Malus* x domestica). *Commun. Agric. Appl. Biol. Sci.* **71**, 187-191.

Makarenkov V (2001) T-REX: reconstructing and visualizing phylogenetic trees and reticulation networks. *Bioinformatics* **17**, 664-668.

Manos PS, Fairbrothers DE (1987) Allozyme variation in populations of six Northeastern American red oaks (Fagaceae, *Quercus* Subg *Erythrobalanus*). *Systematic Botany* **12**, 365-373.

Marthick JR (2005) The phylogeography of two eucalypts. Honours thesis, University of Tasmania.

Matziris DI (1994) Genetic variation in the phenology of flowering in Black Pine. *Silvae Genetica* **43**, 321-328.

McGowen MH (2007) Genetic control of reproductive traits in *E. globulus*. PhD thesis, University of Tasmania.

McGowen MH, Wiltshire RJE, Potts BM, Vaillancourt RE (2001) The origin of *Eucalyptus vernicosa*, a unique shrub eucalypt. *Biological Journal of the Linnean Society* **74**, 397-405.

McKenzie GM (1997) The late Quaternary vegetation history of the south-central highlands of Victoria, Australia .1. Sites above 900 m. *Australian Journal of Ecology* **22**, 19-36.

McKinnon GE, Jordan GJ, Vaillancourt RE, Steane DA, Potts BM (2004a) Glacial refugia and reticulate evolution: the case of the Tasmanian eucalypts. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **359**, 275-284.

McKinnon GE, Potts BM, Steane DA, Vaillancourt RE (2005) Population and phylogenetic analysis of the cinnamoyl coA reductase gene in *Eucalyptus globulus* (Myrtaceae). *Australian Journal of Botany* **53**, 827-838.

McKinnon GE, Steane DA, Potts BM, Vaillancourt RE (1999) Incongruence between chloroplast and species phylogenies in *Eucalyptus* subgenus *Monocalyptus* (Myrtaceae). *American Journal of Botany* **86**, 1038-1046.

McKinnon GE, Vaillancourt RE, Jackson HD, Potts BM (2001a) Chloroplast sharing in the Tasmanian eucalypts. *Evolution* **55**, 703-711.

McKinnon GE, Vaillancourt RE, Steane DA, Potts BM (2004b) The rare silver gum, *Eucalyptus cordata*, is leaving its trace in the organellar gene pool of *Eucalyptus globulus*. *Molecular Ecology* **13**, 3751-3762.

McKinnon GE, Vaillancourt RE, Tilyard PA, Potts BM (2001b) Maternal inheritance of the chloroplast genome in *Eucalyptus globulus* and interspecific hybrids. *Genome* **44**, 831-835.

Mehlenbacher SA, Voordeckers AM (1991) Relationship of flowering time, rate of seed germination, and time of leaf budbreak and usefulness in selecting for late-

flowering apples. *Journal of the American Society for Horticultural Science* **116**, 565-568.

Meilan R (1997) Floral induction in woody angiosperms. *New Forests* **14**, 179-202.

Menzel A, Sparks TH, Estrella N, Koch E, Aasa A, Ahas R, Alm-Kubler K, Bissolli P, Braslavska O, Briede A, Chmielewski FM, Crepinsek Z, Curnel Y, Dahl A, Defila C, Donnelly A, Filella Y, Jatcza K, Mage F, Mestre A, Nordli O, Penuelas J, Pirinen P, Remisova V, Scheifinger H, Striz M, Susnik A, Van Vliet AJH, Wielgolaski FE, Zach S, Zust A (2006) European phenological response to climate change matches the warming pattern. *Global Change Biology* **12**, 1969-1976.

Michalakis Y, Excoffier L (1996) A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics* **142**, 1061-1064.

Missiaggia AA, Piacezzi AL, Grattapaglia D (2005) Genetic mapping of *Eefl*, a major effect QTL for early flowering in *Eucalyptus grandis*. *Tree Genetics & Genomes* **1**, 79-84.

Moncur MW (1992) Effect of low temperature on floral induction of *Eucalyptus lansdowneana* F. Muell and J. Brown ssp. *lansdowneana*. *Australian Journal of Botany* **40**, 157-167.

Moncur MW, Hand FC, Ramsden NG (1994) 'Environmental and cultural effects on flowering and seed production of plantation grown *Eucalyptus nitens*.' Division of Forestry, CSIRO, Canberra.

Moncur MW, Hasan O (1994) Floral induction in *Eucalyptus nitens*. *Tree Physiology* **14**, 1303-1312.

Mouradov A, Cremer F, Coupland G (2002) Control of flowering time: Interacting pathways as a basis for diversity. *Plant Cell* **14**, S111-S130.

Mouradov A, Glassick T, Hamdorf B, Murphy L, Fowler B, Maria S, Teasdale RD (1998) *NEEDLY*, a *Pinus radiata* ortholog of *FLORICAULA/LEAFY* genes, expressed in both reproductive and vegetative meristems. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 6537-6542.

Muneri A, Raymond CA (2000) Genetic parameters and genotype-by-environment interactions for basic density, pilodyn penetration and stem diameter in *Eucalyptus globulus*. *Forest Genetics* **7**, 317-328.

Nei M (1972) Genetic distance between populations. *American Naturalist* **106**, 283-&.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**, 583-890.

Nelson EC (1981) Phytogeography of Southern Australia. In 'Ecological Biogeography of Australia'. (Ed. A Keast) pp. 735-757. (W. Junk: The Hague)

Nesbitt KA, Potts BM, Vaillancourt RE, West AK, Reid JB (1995) Partitioning and distribution of RAPD variation in a forest tree species, *Eucalyptus globulus* (Myrtaceae). *Heredity* **74**, 628-637.

Nikkanen T (2001) Reproductive phenology in a Norway spruce seed orchard. *Silva Fennica* **35**, 39–53.

Nishikawa F, Endo T, Shimada T, Fujii H, Shimizu T, Omura M, Ikoma Y (2007) Increased *CiFT* abundance in the stem correlates with floral induction by low temperature in Satsuma mandarin (*Citrus unshiu* Marc.). *Journal of Experimental Botany* **58**, 3915-3927.

O'Reilly-Wapstra JM, McArthur C, Potts BM (2002) Genetic variation in resistance of *Eucalyptus globulus* to marsupial browsers. *Oecologia* **130**, 289-296.

O'Reilly-Wapstra JM, McArthur C, Potts BM (2004) Linking plant genotype, plant defensive chemistry and mammal browsing in a *Eucalyptus* species. *Functional Ecology* **18**, 677-684.

O'Reilly-Wapstra JM, Potts BM, McArthur C, Davies NW, Tilyard P (2005) Inheritance of resistance to mammalian herbivores and of plant defensive chemistry in an *Eucalyptus* species. *Journal of Chemical Ecology* **31**, 357-375.

Oline DK (2008) Geographic variation in chloroplast haplotypes in the California red fir-noble fir species complex and the status of Shasta red fir. *Canadian Journal of Forest Research* **38**, 2705-2710.

Ozga JA, Yu J, Reinecke DM (2003) Pollination-, development-, and auxin-specific regulation of gibberellin 3 beta-hydroxylase gene expression in pea fruit and seeds. *Plant Physiology* **131**, 1137-1146.

Paetkau D, Strobeck C (1995) The molecular basis and evolutionary history of a microsatellite null allele in bears. *Molecular Ecology* **4**, 519-520.

Parenicova L, de Folter S, Kieffer M, Horner DS, Favalli C, Busscher J, Cook HE, Ingram RM, Kater MM, Davies B, Angenent GC, Colombo L (2003) Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in *Arabidopsis*: New openings to the MADS world. *Plant Cell* **15**, 1538-1551.

Passioura JA, Ash JE (1993) Phenotypic, genetic and ecological variation in the *Eucalyptus saligna-E. botryoides* complex. *Australian Journal of Botany* **41**, 393-412.

Patterson B, Gore P, Potts BM, Vaillancourt RE (2004) Advances in pollination techniques for large-scale seed production in *Eucalyptus globulus*. *Australian Journal of Botany* **52**, 781-788.

Payn KG, Dvorak WS, Janse BJH, Myburg AA (2008) Microsatellite diversity and genetic structure of the commercially important tropical tree species *Eucalyptus urophylla*, endemic to seven islands in eastern Indonesia. *Tree Genetics & Genomes* **4**, 519-530.

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**, 288-295.

Pemberton JM, Slate J, Bancroft DR, Barrett JA (1995) Nonamplifying alleles at microsatellite Loci - a caution for parentage and population studies. *Molecular Ecology* **4**, 249-252.

Pena L, Martin-Trillo M, Juarez J, Pina JA, Navarro L, Martinez-Zapater JM (2001) Constitutive expression of *Arabidopsis LEAFY* or *APETALA1* genes in citrus reduces their generation time. *Nature Biotechnology* **19**, 263-267.

Perez-Lopez D, Ribas F, Moriana A, Rapoport HF, De Juan A (2008) Influence of temperature on the growth and development of olive (*Olea europaea* L.) trees. *Journal of Horticultural Science & Biotechnology* **83**, 171-176.

Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research* **29**.

Pharis RP, Webber JE, Ross SD (1987) The promotion of flowering in forest trees by gibberellin-A4/7 and cultural treatments - a review of the possible mechanisms. *Forest Ecology and Management* **19**, 65-84.

Pillitteri LJ, Lovatt CJ, Walling LL (2004a) Isolation and characterization of a *TERMINAL FLOWER* homolog and its correlation with juvenility in citrus. *Plant Physiology* **135**, 1540-1551.

Pillitteri LJ, Lovatt CJ, Walling LL (2004b) Isolation and characterization of *LEAFY* and *APETALA1* homologues from *Citrus sinensis* L. Osbeck 'Washington'. *Journal of the American Society for Horticultural Science* **129**, 846-856.

Poethig RS (2003) Phase change and the regulation of developmental timing in plants. *Science* **301**, 334-336.

Potts BM, Barbour R, Hingston A (2001) 'The risk of genetic pollution from farm forestry using eucalypt species and hybrids.' Rural Industries Research and Development, Joint Venture Agroforestry Program.

Potts BM, Barbour RC, Hingston AB, Vaillancourt RE (2003) Genetic pollution of native eucalypt gene pools - identifying the risks. *Australian Journal of Botany* **51**, 1-25.

Potts BM, Jordan GJ (1994) The spatial pattern and scale of variation in *Eucalyptus globulus* ssp. *globulus* - variation in seedling abnormalities and early growth. *Australian Journal of Botany* **42**, 471-492.

Potts BM, Vaillancourt RE, Jordan GJ, Dutkowski GW, Costa e Silva J, McKinnon GE, Steane DA, Volker PW, Lopez GA, Apiolaza LA, Li Y, Marques C, Borralho NMG (2004) Exploration of the *Eucalyptus globulus* gene pool. In '*Eucalyptus* in a changing world'. Aveiro, Portugal. (Eds NMG Borralho, JS Pereira, C Marques, J Coutinho, M Madeira and M Tomé) pp. 46-61. (RAIZ, Instituto Investigação de Floresta e Papel)

Potts BM, Volker PW, Dungey HS (1992) Barriers to the production of interspecific hybrids in *Eucalyptus*. In 'Mass Production Technology for Genetically Improved Fast Growing Forest Tree Species (AFOCEL-IUFRO Symposium 1992)'. Bordeaux pp. 193-204. (Association Foret Cellulose)

Potts BM, Wiltshire RJE (1997) Eucalypt genetics and genecology. In 'Eucalypt Ecology: Individuals to Ecosystems'. (Eds J Williams and J Woinarski) pp. 56-91. (Cambridge University Press: Cambridge)

Primack RB (1987) Relationship among flowers, fruits and seeds. *Annual Review of Ecology and Systematics* **18**, 409–430.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.

Pryor LD (1976) 'Biology of Eucalypts.' (Edward Arnold, London.)

Pryor LD, Johnson LAS (1971) 'A Classification of the Eucalypts.' (Australian National University Press: Canberra)

Putterill J, Laurie R, Macknight R (2004) It's time to flower: the genetic control of flowering time. *Bioessays* **26**, 363-373.

Ramos A, Perez-Solis E, Ibanez C, Casado R, Collada C, Gomez L, Aragoncillo C, Allona I (2005) Winter disruption of the circadian clock in chestnut. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 7037-7042.

Rathbone DA, McKinnon GE, Potts BM, Steane DA, Vaillancourt RE (2007) Microsatellite and cpDNA variation in island and mainland populations of a regionally rare eucalypt, *Eucalyptus perriniana* (Myrtaceae). *Australian Journal of Botany* **55**, 513-520.

Rattigan K, Hill SJ (1986) Relationship between temperature and flowering in almond. *Australian Journal of Experimental Agriculture* **26**, 399-404.

Rattigan K, Hill SJ (1988) Relationship between temperature and flowering in almond: effect of location. *Australian Journal of Experimental Agriculture* **27**, 905-908.

Reed DD, Jones EA, Tome M, Araujo MC (2003) Models of potential height and diameter for *Eucalyptus globulus* in Portugal. *Forest Ecology and Management* **172**, 191-198.

Rieseberg LH, Beckstrom-Sternberg SM, Liston A, Arias DM (1991) Phylogenetic and systematic inferences from chloroplast DNA and isozyme variation in *Helianthus* sect *Helianthus* (Asteraceae). *Systematic Botany* **16**, 50-76.

Rieseberg LH, Wendel JF (1993) Introgression and its consequences in plants. In 'Hybrid zones and the evolutionary process'. (Ed. RG Harrison) pp. 70-114. (Oxford University Press: New York Oxford)

Rodriguez-Rajo FJ, Mendez J, Jato V (2005) Factors affecting pollination ecology of Quercus anemophilous species in north-west Spain. *Botanical Journal of the Linnean Society* **149**, 283-297.

Rose TM, Schultz ER, Henikoff JG, Pietrokovski S, McCallum CM, Henikoff S (1998) Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. *Nucleic Acids Research* **26**, 1628-1635.

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes* **4**, 137-138.

Rottmann WH, Meilan R, Sheppard LA, Brunner AM, Skinner JS, Ma CP, Cheng SP, Jouanin L, Pilate G, Strauss SH (2000) Diverse effects of overexpression of *LEAFY* and *PTLF*, a poplar (*Populus*) homolog of *LEAFY/FLORICAULA*, in transgenic poplar and Arabidopsis. *Plant Journal* **22**, 235-245.

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. In 'Bioinformatics Methods and Protocols: Methods in Molecular Biology.' (Eds S Krawetz and S Misener) pp. 365-386. (Humana Press: Totowa, NJ)

Rutledge R, Regan S, Nicolas O, Fobert P, Cote C, Bosnich W, Kauffeldt C, Sunohara G, Seguin A, Stewart D (1998) Characterization of an *AGAMOUS* homologue from the conifer black spruce (*Picea mariana*) that produces floral homeotic conversions when expressed in *Arabidopsis*. *Plant Journal* **15**, 625-634.

Ryman N, Leimar O (2009) G_{ST} is still a useful measure of genetic differentiation - a comment on Jost's *D*. *Molecular Ecology* **18**, 2084-2087.

Saitou N, Nei M (1987) The Neighbor-Joining method - a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**, 406-425.

Sane AP, Tripathi SK, Nath P (2007) Petal abscission in rose (*Rosa bourboniana* var Gruss an Teplitz) is associated with the enhanced expression of an alpha expansin gene, *RbEXPA1*. *Plant Science* **172**, 481-487.

Silva C, Garcia-Mas J, Sanchez AM, Arus P, Oliveira M (2005) Looking into flowering time in almond (*Prunus dulcis* (Mill) D. A. Webb): the candidate gene approach. *Theoretical and Applied Genetics* **110**, 959-968.

Skabo S, Vaillancourt RE, Potts BM (1998) Fine-scale genetic structure of *Eucalyptus globulus* ssp. *globulus* forest revealed by RAPDs. *Australian Journal of Botany* **46**, 583-594.

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**, 457-462.

Slee A, Brooker M, Duffy S, West J (2006) 'EUCLID: eucalypts of Australia.' (CSIRO Publishing: Collingwood Australia)

Smith S, Hughes J, Wardell-Johnson G (2003) High population differentiation and extensive clonality in a rare mallee eucalypt: *Eucalyptus curtisii* - Conservation genetics of a rare mallee eucalypt. *Conservation Genetics* **4**, 289-300.

Somerville C, Koornneef M (2002) Timeline - A fortunate choice: the history of *Arabidopsis* as a model plant. *Nature Reviews Genetics* **3**, 883-889.

Southerton SG (2007) Early flowering induction and *Agrobacterium* transformation of the hardwood tree species *Eucalyptus occidentalis*. *Functional Plant Biology* **34**.

Southerton SG, Marshall H, Mouradov A, Teasdale RD (1998a) Eucalypt MADS-box genes expressed in developing flowers. *Plant Physiology* **118**, 365-372.

Southerton SG, Strauss SH, Olive MR, Harcourt RL, Decroocq V, Zhu XM, Llewellyn DJ, Peacock WJ, Dennis ES (1998b) Eucalyptus has a functional equivalent of the *Arabidopsis* floral meristem identity gene *LEAFY*. *Plant Molecular Biology* **37**, 897-910.

Spano D, Cesaraccio C, Duce P, Snyder RL (1999) Phenological stages of natural species and their use as climate indicators. *International Journal of Biometeorology* **42**, 124-133.

Sreekantan L, Clemens J, McKenzie MJ, Lenton JR, Croker SJ, Jameson PE (2004) Flowering genes in *Metrosideros* fit a broad herbaceous model encompassing *Arabidopsis* and *Antirrhinum*. *Physiologia Plantarum* **121**, 163-173.

Steane DA, Byrne M, Vaillancourt RE, Potts BM (1998) Chloroplast DNA polymorphism signals complex interspecific interactions in *Eucalyptus* (Myrtaceae). *Australian Systematic Botany* **11**, 25-40.

Steane DA, Conod N, Jones RC, Vaillancourt RE, Potts BM (2006) A comparative analysis of population structure of a forest tree, *Eucalyptus globulus* (Myrtaceae), using microsatellite markers and quantitative traits. *Tree Genetics & Genomes* **2**, 30-38.

Steane DA, Vaillancourt RE, Russell J, Powell W, Marshall D, Potts BM (2001) Development and characterisation of microsatellite loci in *Eucalyptus globulus* (Myrtaceae). *Silvae Genetica* **50**, 89-91.

Stebbins G (1989) Plant speciation and the founder principle. In 'Genetics, Speciation and the Founder Principle'. (Eds L Giddings, K Kaneshiro and W Anderson) pp. 113-125. (Oxford University Press: New York)

Swofford DL (2003) 'PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4.' (Sinauer Associates: Sunderland, Massachusetts)

Sykes MT, Prentice IC (1996) Climate change, tree species distributions and forest dynamics: A case study in the mixed conifer/northern hardwoods zone of Northern Europe. *Climatic Change* **34**, 161-177.

Tan FC, Swain SM (2007) Functional characterization of *AP3*, *SOC1* and *WUS* homologues from citrus (*Citrus sinensis*). *Physiologia Plantarum* **131**, 481-495.

Taylor JE, Whitelaw CA (2001) Signals in abscission. *New Phytologist* **151**, 323-339.

Thompson R, Clark RM (2006) Spatio-temporal modelling and assessment of within-species phenological variability using thermal time methods. *International Journal of Biometeorology* **50**, 312-322.

Tovar-Sanchez E, Oyama K (2004) Natural hybridization and hybrid zones between *Quercus crassifolia* and *Quercus crassipes* (Fagaceae) in Mexico: Morphological and molecular evidence. *American Journal of Botany* **91**, 1352-1363.

Tripiana V, Bourgeois M, Verhaegen D, Vigneron P, Bouvet JM (2007) Combining microsatellites, growth, and adaptive traits for managing *in situ* genetic resources of *Eucalyptus urophylla*. *Canadian Journal of Forest Research* **37**, 773-785.

Trudgill DL, Honek A, Li D, Straalen NM (2005) Thermal time - concepts and utility. *Annals of Applied Biology* **146**, 1-14.

Turner C, Wiltshire, R. J. E., Potts, B.M. Vaillancourt, R.E. (2000) Allozyme variation and conservation of the Tasmanian endemics, *Eucalyptus risdonii*, *E. tenuiramis* and *E. coccifera*. *Conservation Genetics* **1**, 209-216.

Vahala J, Keinanen M, Schutzendubel A, Polle A, Kangasjarvi J (2003) Differential effects of elevated ozone on two hybrid aspen genotypes predisposed to chronic ozone fumigation. Role of ethylene and salicylic acid. *Plant Physiology* **132**, 196-205.

Vahala T, Oxelman B, von Arnold S (2001) Two *APETALA2*-like genes of *Picea abies* are differentially expressed during development. *Journal of Experimental Botany* **52**, 1111-1115.

Vaillancourt RE, Boomsma DB, Nicolle D (2001) A disjunct population of *Eucalyptus globulus* ssp. *bicostata* from South Australia. *Transactions of the Royal Society of South Australia* **125**, 65-68.

van Asch M, Tienderen PH, Holleman LJM, Visser ME (2007) Predicting adaptation of phenology in response to climate change, an insect herbivore example. *Global Change Biology* **13**, 1596-1604.

van der Linden CG, Vosman B, Smulders MJM (2002) Cloning and characterization of four apple MADS box genes isolated from vegetative tissue. *Journal of Experimental Botany* **53**, 1025-1036.

Visser T, Schaap AA (1967) Pre-selection for juvenile period, flowering and picking time in apple seedlings. *Euphytica* **16**, 109–121.

Wada M, Cao QF, Kotoda N, Soejima J, Masuda T (2002) Apple has two orthologues of *FLORICAULA/LEAFY* involved in flowering. *Plant Molecular Biology* **49**, 567-577.

Wareing PF, Saunders PF (1971) Hormones and dormancy. *Annual Review of Plant Physiology* **22**, 261-288.

Watson JM, Brill EM (2004) *Eucalyptus grandis* has at least two functional *SOC1*-like floral activator genes. *Functional Plant Biology* **31**, 225-234.

Weigel D, Nilsson O (1995) A developmental switch sufficient for flower initiation in diverse plants. *Nature* **377**, 495-500.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358-1370.

Whitham TG, Bailey JK, Schweitzer JA, Shuster SM, Bangert RK, Leroy CJ, Lonsdorf EV, Allan GJ, DiFazio SP, Potts BM, Fischer DG, Gehring CA, Lindroth RL, Marks JC, Hart SC, Wimp GM, Wooley SC (2006) A framework for community and ecosystem genetics: from genes to ecosystems. *Nature Reviews Genetics* **7**, 510-523.

Whitham TG, DiFazio SP, Schweitzer JA, Shuster SM, Allan GJ, Bailey JK, Woolbright SA (2008) Perspective - Extending genomics to natural communities and ecosystems. *Science* **320**, 492-495.

Whitham TG, Gehring CA, Evans LM, LeRoy CJ, Bangert RK, Schweitzer JA, Allan GJ, Barbour RC, Fischer DG, Potts BM, Bailey JK (2009) A community and ecosystem genetics approach to conservation biology and management. In 'Molecular Approaches in Natural Resource Conservation'. (Eds JA DeWoody, JW Bickham, C Michler, K Nichols, OE Rhodes and K Woeste) pp. in press. (Cambridge University Press)

Whitham TG, Martinsen GD, Floate KD, Dungey HS, Potts BM, Keim P (1999) Plant hybrid zones affect biodiversity: Tools for a genetic-based understanding of community structure. *Ecology* **80**, 416-428.

Whittock SP, Apiolaza LA, Kelly CM, Potts BM (2003) Genetic control of coppice and lignotuber development in *Eucalyptus globulus*. *Australian Journal of Botany* **51**, 57-67.

Wigge PA, Kim MC, Jaeger KE, Busch W, Schmid M, Lohmann JU, Weigel D (2005) Integration of spatial and temporal information during floral induction in *Arabidopsis*. *Science* **309**, 1056-1059.

Wilkie JD, Sedgley M, Olesen T (2008) Regulation of floral initiation in horticultural trees. *Journal of Experimental Botany*.

Williams DR, Potts BM, Neilsen WA, Joyce KR (2006) The effect of tree spacing on the production of flowers in *Eucalyptus nitens*. *Australian Forestry* **69**, 299-304.

Williams KJ, Potts BM (1996) The natural distribution of *Eucalyptus* species in Tasmania. *Tasforests* **8**, 39-165.

Xu Y, Zhang L, Xie H, Zhang YQ, Oliveira MM, Ma RC (2008) Expression analysis and genetic mapping of three *SEPALLATA*-like genes from peach (*Prunus persica* (L.) Batsch). *Tree Genetics & Genomes* **4**, 693-703.

Yamane H, Kashiwa Y, Ooka T, Tao R, Yonemori K (2008) Suppression subtractive hybridization and differential screening reveals endodormancy-associated expression of an *SVP/AGL24*-type MADS-box gene in lateral vegetative buds of japanese apricot. *Journal of the American Society for Horticultural Science* **133**, 708-716.

Yan LL, Loukoianov A, Blechl A, Tranquilli G, Ramakrishna W, SanMiguel P, Bennetzen JL, Echenique V, Dubcovsky J (2004) The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization. *Science* **303**, 1640-1644.

Yao JL, Dong YH, Kvarnheden A, Morris B (1999) Seven MADS-box genes in apple are expressed in different parts of the fruit. *Journal of the American Society for Horticultural Science* **124**, 8-13.

Yao JL, Dong YH, Morris BAM (2001) Parthenocarpic apple fruit production conferred by transposon insertion mutations in a MADS-box transcription factor. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 1306-1311.

Yatabe Y, Kane NC, Scotti-Saintagne C, Rieseberg LH (2007) Rampant gene exchange across a strong reproductive barrier between the annual sunflowers, *Helianthus annuus* and *H. petiolaris*. *Genetics* **175**, 1883-1893.

Zhang L, Xu Y, Ma RC (2008) Molecular cloning, identification, and chromosomal localization of two MADS box genes in peach (*Prunus persica*). *Journal of Genetics and Genomics* **35**, 365-372.

Zhou S, Su H, Wang L, Zhang X, Shu H (2006) Cloning and transformation of *APETALA2* homologues genes from apple. *Acta Horticulturae Sinica* **33**, 239-243.